

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:43:31 ; Search time 41 Seconds  
(without alignments)  
940.800 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNHGPDTEAEEDFDFMTV.....VTDEIVXEFMTPRKLSDFDF 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619-526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 2101   | 99.3        | 471    | 1 A41706 | tryptophan-tRNA li |
| 2          | 2028.5 | 95.9        | 475    | 1 YWBO   | tryptophan-tRNA li |
| 3          | 1946.5 | 92.0        | 475    | 1 YWRBPR | tryptophan-tRNA li |
| 4          | 1938   | 91.6        | 481    | 2 S50053 | tryptophan-tRNA li |
| 5          | 1210   | 57.2        | 395    | 2 S58157 | hypothetical prote |
| 6          | 1163   | 55.0        | 432    | 2 S51901 | tryptophan-tRNA li |
| 7          | 967    | 42.9        | 386    | 2 C90190 | tryptophan-tRNA li |
| 8          | 903    | 37.9        | 385    | 2 C75020 | tryptophanyl-tRNA  |
| 9          | 626.5  | 29.6        | 301    | 2 G71206 | tryptophan-tRNA li |
| 10         | 534.5  | 25.3        | 380    | 2 G84373 | tryptophanyl-tRNA  |
| 11         | 409.5  | 19.4        | 370    | 2 F64476 | tryptophan-tRNA li |
| 12         | 397.5  | 18.8        | 364    | 2 B69131 | tryptophan-tRNA li |
| 13         | 386    | 18.2        | 334    | 2 T43806 | tryptophan-tRNA li |
| 14         | 370.5  | 17.5        | 420    | 2 B69461 | tryptophanyl-tRNA  |
| 15         | 353    | 16.7        | 374    | 2 D72477 | probable tryptopha |
| 16         | 263.5  | 12.7        | 533    | 2 F84371 | tryptophanyl-tRNA  |
| 17         | 192    | 9.1         | 323    | 2 H69346 | tyrosyl-tRNA synth |
| 18         | 188    | 8.9         | 364    | 2 F72512 | probable tyrosyl-t |
| 19         | 185.5  | 8.8         | 341    | 2 D95260 | tryptophanyl-tRNA  |
| 20         | 185.5  | 8.8         | 341    | 2 G98125 | tryptophan-tRNA li |
| 21         | 174.5  | 8.2         | 341    | 2 B86633 | tryptophan-tRNA li |
| 22         | 172    | 8.1         | 334    | 2 A45993 | tyrosine-tRNA liga |
| 23         | 170.5  | 8.1         | 366    | 2 S75410 | tryptophan-tRNA li |
| 24         | 169    | 8.0         | 335    | 2 H70385 | tryptophan-tRNA li |
| 25         | 167    | 7.9         | 351    | 2 E75438 | tryptophanyl-tRNA  |
| 26         | 166.5  | 7.9         | 346    | 2 B71496 | tryptophan-tRNA li |
| 27         | 164.5  | 7.8         | 335    | 2 A86410 | protein F3M18.22   |
| 28         | 163.5  | 7.7         | 333    | 2 E70100 | tryptophan-tRNA li |
| 29         | 163    | 7.7         | 460    | 2 C84750 | probable tyrosyl-t |

#### ALIGNMENTS

##### RESULT 1

A41706  
tryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: interferon-inducible protein IFP53; peptide-chain release factor homol  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text change 03-Jun-2002  
C:Accession: A41633; A41706; S19246; JNC676; JH0533; S26287  
R:Fleckenner, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11523-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (gamma  
A:Reference number: A41633; MUID:92107982; PMID:1763065  
A:Accession: A41633  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <FILE>  
A:Cross-references: GB:X59892; NID:g33820; PIDN:CAA42545.1; PID:g30821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophanyl-tRNA synthetase expression in human fibroblasts  
A:Reference number: A41706; MUID:92105071; PMID:1761529  
A:Accession: A41706  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:g1846656; PIDN:AAA67324.1; PID:g184657  
R:Buwitt, U.; Flohr, T.; Boettger, E.C.  
EMBO J. 11, 489-496, 1992  
A:Title: Molecular cloning and characterization of an interferon induced human cDNA with  
A:Reference number: S19246; MUID:92164636; PMID:1537332  
A:Accession: S19246  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-423, R, 425-471 <BUW>  
A:Cross-references: EMBL:X62570; NID:g32708; PIDN:CAA44450.1; PID:g32709  
R:Prolova, L.Y.; Grigorova, A.Y.; Sudomocina, M.A.; Kisselev, L.L.  
Gene 128, 217-245, 1993  
A:Title: The human gene encoding tryptophanyl-tRNA synthetase: Interferon-response element  
A:Reference number: JN0676; MUID:93292992; PMID:7695728  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141,182-471 <PROI>  
A:Cross-references: GB:S62837; NID:g37968; GB:X67919; NID:g37969; GB:X67920; N  
4; NID:g37974; GB:X67925; GB:S62855; NID:g37975; GB:X67926; GB:S62856; NID:g37976; GB:X6  
A:Note: The authors translated the codon GGG for residue 55 as Cys and GAG for residue 34  
R:Prolova, L.Y.; Sudomocina, M.A.; Grigorova, A.Y.; Zinovieva, O.L.; Kisselev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human trypt  
A:Reference number: JH0533; MUID:92112058; PMID:1765274  
A:Accession: JH0533  
A:Molecule type: mRNA

30 159.5 7.5 337 2 F71300  
31 158.5 7.5 375 2 B75072  
32 158 7.5 346 2 C81654  
33 154.5 7.3 102 2 T44994  
34 152 7.2 344 2 H86590  
35 152 7.2 344 2 C72034  
36 149.5 7.1 327 2 C84374  
37 149.5 7.1 408 2 T03741  
38 146 6.9 339 2 E64676  
39 146 6.9 375 2 F71093  
40 145 6.9 337 2 A11086  
41 138 6.5 319 2 H69102  
42 135.5 6.4 328 2 C72370  
43 135 6.4 365 2 E82052  
44 132.5 6.3 343 2 S73024  
45 132 6.2 328 1 YWBSF

A:Residues: 1-212,'GD',215-471 <PRO>  
 A:Cross-references: GB:M61715; NID:G340367; PIDN:AAA61298.1; PID:G340368  
 A:Experimental source: fibroblast

C:Genetics: GDB:WARS; IFF53  
 A:Gene: GDB:WARS; IFF53  
 A:Cross-references: GDB:119632; OMIM:191050  
 A:Map position: 14Q23-14Q31  
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:19-64/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2101; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 1.7e-163;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SNHGDATEAEEDFVDPWTQTSAAKIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 DB 71 SNHGDATEAEEDFVDPWTQTSAAKIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
 QY 61 FLRRGIFFSHRDMQVLDAYENKPFYLTGRPSSEAMHVGHLIPFIETKWLQDVFNVP 120  
 DB 131 FLRRGIFFSHRDMQVLDAYENKPFYLTGRPSSEAMHVGHLIPFIETKWLQDVFNVP 190  
 QY 121 LVIQMTDDKYLWKDLTLQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 DB 191 LVIQMTDDKYLWKDLTLQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250  
 QY 181 VKIKQKHTFNVKGIFFGFTSDCKIGKISFPALQAPSFNSFPQIFRDRDIIQCLIPCA 240  
 DB 251 VKIKQKHTFNVKGIFFGFTSDCKIGKISFPALQAPSFNSFPQIFRDRDIIQCLIPCA 310  
 QY 241 IDQPYFRMTRDVAPRIGYKPEALLHSTFFPALQAGTQKMSASDPNSSIFLDTAKQIKT 300  
 DB 311 IDQPYFRMTRDVAPRIGYKPEALLHSTFFPALQAGTQKMSASDPNSSIFLDTAKQIKT 370  
 QY 301 KVNKHAFFSGRDTIEHRQFGGNCVDVSVFMYLTFFLEDDCKLEQIRKDYTSGLTSEL 360  
 DB 371 KVNKHAFFSGRDTIEHRQFGGNCVDVSVFMYLTFFLEDDCKLEQIRKDYTSGLTSEL 430  
 QY 361 KKALIEVLQPLIAEHOARKEVTDIVKEFMTPRKLSDFQ 401  
 DB 431 KKALIEVLQPLIAEHOARKEVTDIVKEFMTPRKLSDFQ 471

## RESULT 2

WYRBP

tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - bovine  
 N:Alternate names: tryptophanyl-tRNA synthetase  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text\_change 03-Jun-2002  
 C:Accession: A40279; JN0354; S10460; S14540  
 R:Lee, C.C.; Craig, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequence  
 A:Reference number: A40279; MUID:91329348; PMID:1907847  
 A:Accession: A40279  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <GAP>  
 C:Cross-references: GDB:M74074; EMBL:X53918; NID:G163798; PIDN:AAA30799.1; PID:G163799  
 A:Experimental source: pancreas  
 A:Note: The authors translated the codon CTG for residue 347 as Ala and CAG for residue  
 A:Note: part of this sequence was confirmed by protein sequencing  
 A:Zargava, T.A.; Kovaleva, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.  
 A:Biochem. Khim. 15, 1307-1311, 1989  
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca  
 A:Reference number: JN0354; MUID:90211408; PMID:2631684  
 A:Accession: JN0354  
 A:Molecule type: protein  
 A:Residues: 1-124;282-287,'N',288,'F',293-294,'R',336-353,423-441,443-449  
 A:Experimental source: liver  
 A:Note: this paper is in Russian  
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:24-69/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 95.9%; Score 2028.5; DB 1; Length 475;  
 Best Local Similarity 95.8%; Pred. No. 1.4e-157;  
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;  
 QY 1 SNHGDATEAEEDFVDPWTQTSAAKIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 DB 76 SEGLDATEADEDFVDPWTQTSAAKIDYDKLIVRFGSSKIDKELINRIERATGQRP 135  
 QY 61 FLRRGIFFSHRDMQVLDAYENKPFYLTGRPSSEAMHVGHLIPFIETKWLQDVFNVP 120  
 DB 136 FLRRGIFFSHRDMQVLDAYENKPFYLTGRPSSEAMHVGHLIPFIETKWLQDVFNVP 195  
 QY 121 LVIQMTDDKYLWKDLTLQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 DB 196 LVIQMTDDKYLWKDLTLQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 254  
 QY 181 VKIKQKHTFNVKGIFFGFTSDCKIGKISFPALQAPSFNSFPQIFRDRDIIQCLIPCA 240  
 DB 255 VKIKQKHTFNVKGIFFGFTSDCKIGKISFPALQAPSFNSFPQIFRDRDIIQCLIPCA 314  
 QY 241 IDQPYFRMTRDVAPRIGYKPEALLHSTFFPALQAGTQKMSASDPNSSIFLDTAKQIKT 300  
 DB 315 IDQPYFRMTRDVAPRIGYKPEALLHSTFFPALQAGTQKMSASDPNSSIFLDTAKQIKT 374  
 QY 301 KVNKHAFFSGRDTIEHRQFGGNCVDVSVFMYLTFFLEDDCKLEQIRKDYTSGLTSEL 360  
 DB 375 KVNKHAFFSGRDTIEHRQFGGNCVDVSVFMYLTFFLEDDCKLEQIRKDYTSGLTSEL 434  
 QY 361 KKALIEVLQPLIAEHOARKEVTDIVKEFMTPRKLSDFQ 401  
 DB 435 KKALIEVLQPLIAEHOARKEVTDIVKEFMTPRKLSDFQ 475

## RESULT 3

WYRBP

tryptophan-tRNA ligase (EC 6.1.1.2) [validated] - rabbit  
 N:Alternate names: tryptophanyl-tRNA synthetase  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 30-Sep-1992 #sequence revision 13-Feb-1998 #text\_change 03-Jun-2002  
 C:Accession: A35904; S37396  
 R:Lee, C.C.; Craig, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequence  
 A:Reference number: A35904; MUID:90239043; PMID:2185472  
 A:Accession: A35904  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <LEE>  
 C:Cross-references: GDB:M33460  
 R:Frolova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Drugeon, G.; McCaughan, K.K.;  
 EXBO J. 12, 4013-4019, 1993  
 A:Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase are  
 A:Reference number: S37396; MUID:94009008; PMID:8404867  
 A:Accession: S37396  
 A:Molecule type: mRNA  
 A:Residues: 166-177 <PRO>  
 C:Genetics:  
 A:Gene: WRS  
 C:Complex: homodimer [validated, MUID:94009008]  
 C:Function:  
 A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent formati  
 A:Note: mammalian WRS (tryptophanyl-tRNA synthetases) and eRF (polypeptide chain release  
 C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
 C:Keywords: aminoacyl-tRNA synthetase; ATP; homodimer; ligase; metalloprotein; protein b  
 F:23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>  
 F:174-177/Region: ATP-binding motif (HXHG)

Query Match 92.0%; Score 1946.5; DB 1; Length 475;  
 Best Local Similarity 90.8%; Pred. No. 6.9e-151;  
 Matches 364; Conservative 2; Mismatches 15; Indels 1; Gaps 1;



```

124 DDEKFLFKQVSLSDQCFARENAKDIIVAGDPKKTFFIMNSTYVG--GAFYQNVVRIA 181
186 KHVTFNQVKGIFGFTSDICIGKISPAIAQAAPSFNSPQIFRDRDIOCLIPCAIDQDP 245
182 KCITANQSKACGFTSDSISGKIHFAISIAQAAPSFNSPQIFRDRDIOCLIPCAIDQDP 241
246 YFRMTDVAIPRIGYKPKPALLHSTFFPALQCAQTKMSASDPNSIFLDTAKQIKTKVNKH 305
242 YRLTFEDVSLRLKFKPKPALLHSTFFPALQCAQTKMSASDPNSIFLDTAKQIKTKVNKH 301
306 AFSGGDTTIEHRQFGNCDVDSFVWYLFTEDEDDKLEQIRKDYTSAGMLTGELKKALI 365
302 AFSGGGATTEIHRKQGNPDVAVYQFLSFFLDDDDKLEQIRKDYTSAGMLTGELKKALI 361
366 EYLOPLIAHQARRKEVTDEIVKEKMT-PRKLSF 398
362 KULQOFVDFQARSKVDEATLDMFMGGRKLEW 395

RESULT 6
S51901
  Cryptophan-tRNA ligase (EC 6.1.1.2) [similarity] - yeast (Saccharomyces cerevisiae)
  N:Altername names: protein HR432; protein O092; tryptophanyl-tRNA synthetase
  C:Species: Saccharomyces cerevisiae
  C:Date: 05-May-1995 #sequence revision 03-Aug-1995 #text_change 03-Jun-2002
  C:Accession: S51901; S59177; S66793
  R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
  submitted to the EMBL Data Library, January 1995
  A:Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including
    and a Delta.
  A:Reference number: S51848
  A:Accession: S51901
  A:Molecule type: DNA
  A:Residues: 1-432 <VAV>
  A:Cross-references: EMBL:Z48149; NID:G663234; PIDN:CAA88164.1; PID:G663256
  R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
  Yeast 11, 1069-1075, 1995
  A:Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the
    a delta element.
  A:Reference number: S59156; MUID:96076631; PMID:7502582
  A:Accession: S59177
  A:Status: nucleic acid sequence not shown; translation not shown
  A:Molecule type: DNA
  A:Residues: 1-432 <VAV>
  A:Cross-references: EMBL:Z48149; NID:G663234; PIDN:CAA88164.1; PID:G663256
  A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
  R:Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.
  submitted to the Protein Sequence Database, July 1996
  A:Reference number: S66791
  A:Accession: S66793
  A:Molecule type: DNA
  A:Residues: 1-432 <DUR>
  A:Cross-references: EMBL:Z74839; NID:G141947; PIDN:CAA99110.1; PID:G1419948; GSPDB:GN00
  A:Experimental source: strain S288C
  C:Genetics:
  A:Gene: SGD:WRS1; WRS1; MIPS:YOL097C
  A:Map position: 15L
  A:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
  C:Keywords: ligase
  F:117-120/Region: ATP-binding motif (HXGH)

  Query Match 55.0%; Score 1163; DB 2; Length 432;
  Best Local Similarity 54.8%; Pred. No. 5e-87;
  Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

  QY 7 ATEAEDFVDPWT-----QTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQPHHFF 61
  DB 19 STDVKEQVTPWDVEGVDEQGAQNIIDYDKLIVRFGSSKIDKELINRIERATGQPHHFF 78
  QY 62 LRGGIFSFHRDKNQVLDAYENKPFYLTGRGSSSAMHVGHLIPFIETKLODYVNTFLVQMT 121
  DB 117-120/Region: ATP-binding motif (HXGH)

```

```

DB 79 LRKGLFFSRDFTKILDYVEQKPFYLTGRGSSSDMHZGHMIPFVFTKWLQEVFVPL 138
QY 122 VIQMTDDEKYLWK-DLTLCAYGDAVENAKDIITACGPOINKTIFISDLDYMGSSGFYKN 180
DB 139 VIETDDEKFLFKHUKITINDVKNFAENAKDIIVAGDPKKTFFIMNSTYVG--GAFYET 196
QY 181 VKIQKHVTFNQVKGIFGFTSDICIGKISPAIAQAAPSFNSPQIFRDRDIOCLIPCA 240
DB 197 VWRVSQITGSTAKAVFGNDSICIGKEHFAISIAQAAPSFNSPQIFRDRDIOCLIPCA 256
QY 241 IDDPYFRMTDVAIPRIGYKPKPALLHSTFFPALQCAQTKMSASDPNSIFLDTAKQIKT 300
DB 257 IDDPYFRMTDVAIPRIGYKPKPALLHSTFFPALQCAQTKMSASDPNSIFLDTAKQIKT 316
QY 301 KVNKHAFSGRDITIEHRQFGNCDVDSFVWYLFTEDEDDKLEQIRKDYTSAGMLTGEL 360
DB 317 KINKYAFSGGQVADLHRELGGPDVAVYQFLSFFLDDDDKLEQIRKDYTSAGMLTGEL 376
QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEKMT-PRKLS 395
DB 377 KKLCIETLQEFVKAFQERRAQVDEETLQKFMVPHKL 412

RESULT 7
C90190
  tryptophanyl-tRNA synthetase (trps) [imported] - Sulfolobus solfataricus
  C:Species: Sulfolobus solfataricus
  C:Date: 24-May-2001 #sequence revision 24-May-2001 #text_change 15-Jun-2001
  C:Accession: C90190
  R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
    Y. Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.;
    arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
    submitted to GenBank, April 2001
  A:Description: Sulfolobus solfataricus complete genome.
  A:Reference number: A99139
  A:Accession: C90190
  A:Status: preliminary
  A:Molecule type: DNA
  A:Residues: 1-386 <KUR>
  A:CROSS-references: GB:AE006641; NID:G13813608; PIDN:AAK40778.1; GSPDB:GN00155
  C:Genetics:
  C:Gene: trps
  C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

  Query Match 42.9%; Score 907; DB 2; Length 386;
  Best Local Similarity 48.6%; Pred. No. 3.3e-66;
  Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

  QY 8 TEAEDFVDPWTQTSAGK-IDYDKLIVRFGSSKIDKELINRIERATGQPHHFFIRGI 66
  DB 6 TMDPDTVTPWEV-----KGVVDYDKLIVQFGTKITEELKQRIKXNLAGDL-HWMLARNV 59
  QY 67 FFSHRDMQVLDAYENKPFYLTGRGSSAMHVGHLIPFIETKLODYVNTFLVQMT 126
  DB 60 FFSHRDLVDNDEYKSGFFFLYTRAPSL-GHGHIGLIPFIETKLODYVNTFLVQMT 118
  QY 127 DDEKYLWK-DLTLCAYGDAVENAKDIITACGPOINKTIFISDLDYMGSSGFYKNVYKIQ 185
  DB 119 DDEKYLWK-DLTLCAYGDAVENAKDIITACGPOINKTIFISDLDYMGSSGFYKNVYKIQ 175
  QY 186 KHVTFNQVKGIFGFTSDICIGKISPAIAQAAPSFNSPQIFRDRDIOCLIPCAIDQDP 245
  DB 176 KKLTFSEVRATFGDASSNIGLIFYPALQIAPT-----MFEKK---RCLIPAGIDQDP 225
  QY 246 YFRMTDVAIPRIGYKPKPALLHSTFFPALQCAQTKMSASDPNSIFLDTAKQIKTKVNKH 305
  DB 226 YMLQDIDIESLYGKAAQIHSKFLPFLTGPEKSSSNPNPETAIIYVDVDPKTVKIMKY 285
  QY 306 AFSGGDTTIEHRQFGNCDVDSFVWYLFTEDEDDKLEQIRKDYTSAGMLTGELKKAL 364
  DB 286 AFSGGGQFTIELHRKYGNPDVAVYQFLSFFLDDDDKLEQIRKDYTSAGMLTGELKKAL 345
  QY 365 IEVLQPLIAHQARRKEVTDEIVKEKMT-PRKLS 397

```

A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH1921  
C;Superfamily: yeast tyrosine-tRNA ligase  
C;Keywords: aminoacyl-t-RNA synthetase; ligase; protein biosynthesis

Query Match 29.6%; Score 626.5; DB 2; Length 301;  
Best Local Similarity 44.2%; Pred. No. 1.8e-43;  
Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

QY 99 MHNCHLPPFTKMLQDVNTPLVMTDDEKYLNKD-LTLDAQGDAYENAKDIACGF 157  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||  
Db 1 MHGHHIPFFATKLQEKFGVNLVIQTDDDEKFLPKENLTFTDTTRWAYDNLDIIAVGF 60  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 158 DINKTTFISDLUDYGMSSSGFKNWVKIKKHVTFNQVKIGFGFTSDCIKGISFPALQAAP 217  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 61 DPDKTFIFONSEF-----TKIYEMAIPIAKKNIFSMAKAVGFTQSKIIMIFPALQIAP 116  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 218 SFSNSFPQIFRDRTDIOCLICADIQDPYFRMTRDVAPRIGYPKPALLESTFFPALQGAQ 277  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 117 TF-----PERK---RCLPAAIDQPYWRQLORDFAESLGYYKTAALHSKEVPSTLSLS 166  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 278 TKMSASPNSSIFPTHTAKQIKTKVNKAHAFSGGRGTIEHRQFGNCVDVDSFMYLTFEL 337  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 167 GKMSASKPETAIYLTDSZEDVEKKWKFTLTGGRTPLKQRREKGGEPRCKCVFKWELEIF 226  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 338 EDDDKLQRIKDY--TSGLMITGELKKALLEVLQPLAEHQARKEVTDIVKEFWTPR 394  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 227 EDDDK-KLKERYACKNGELTGCECKRYLISKIQEFLKEHQRRRAK-AEKLVERKTYG 283  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 395 KLS 397  
|||:  
Db 284 KIA 286

RESULT 10  
S84373  
tryptophanyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C;Accession: S84373  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
J;Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A54160; MUID:20504483; PMID:11016950  
A;Accession: S84373  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-380 <STO>  
A;Cross-references: GE:AE004437; NID:g1c581646; PIDN:AAG20355.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: trpS2  
C;Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 25.3%; Score 534.5; DB 2; Length 380;  
Best Local Similarity 35.4%; Pred. No. 8e-36;  
Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

QY 10 AEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRRGIFFS 69  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 3 AGNDVTPIAVESDD----LYTEKLIAREGAEUETDQARFP-----DHPTVNRGLFTA 53  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 70 HRDMNQVLDAYENKKKFLLYTGRGPSSSAMVGHILPIFTKWLQDVNFNPZVICTMDDE 129  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 54 GRDVDDELTAEGQS----IVTVGVPSG-PMEIGHAMVVYFAELLQDEFGARVYVPLSDDE 108  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

QY 130 KYLWKDLTLQAYGDAVE-NAKDIACGFDINKTPIF---SDLXY-GMSSGFYKXWVKI 184  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

Db 109 KYWFQDOTPAET-GGYLRANVRDLIAVGDFDELTRIIVDTRADVLPPLATAFNGSV--- 164  
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||

185 QXHTVFNQVKGIFGFTDSCIGKISPPAIAQAPSFNSPFIQIPRDRDIOCLIPCAIDQ 244  
165 -RHATLQNVG-----EPDNVGAFFPAVQAKL---LPQLVHG--EHTLPIAVDQD 213  
245 PYRMTDRVAPRIQYP--KPALLHSTFPFALQQAOTKMSADPNSSIFLTDITAKQIKTV 302  
214 PHRVSRDVAAKARYPVGKPCALLMQF--PSLAG--PGKSSS--AGVSRLTDSPTVREKV 271  
303 NKHAFSGRDTIEHROFGNCDVDSFMYLTFFLEDD--KLQIRKDYTSGLMTGELX 361  
272 RTHAYTGRASVSEHRKGVPAEDWFQLSAFFEPDDAEALRIERYAGDILLSGELK 331  
362 KALIEVLQPLIAHQARKEVTDIVKFTPRKLSFD 399  
332 DLAAADRITEFLAAHQRRALGD--VTEALDAFRLTDD 367

RESULT 11  
F64476  
tryptophan-tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 03-Jun-2002  
C:Accession: F64476  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.E.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8698087  
A:Accession: F64476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-370 <SUL>  
A:Cross-references: GB:U67582; GB:L77117; NID:gl592064; PIDN:AAB99425.1; PID:gl592065; T  
C:Genetics:  
A:Map position: FOR1375885-1376997  
A:Start codon: GTG  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 19.4%; Score 409.5; DB 2; Length 370;  
Best Local Similarity 30.9%; Pred. No. 1.2e-25;  
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
17 PWTVTSSAKGIDYDKLIVRFSGSKIDKELINIERATGQPHFLRRGIFPFSHRDMQV 76  
8 PW--ETPAV--IDYKTMBOFGVKPIVWZGDLKEE-----HHFFRNILGHRDPERI 57  
77 LDAYENKKPFYLYTGRGSPSEAMHVGHLIPFIETKWLQ---DVFNVLV--QMTDDEKYLW 133  
58 VDAIKNKKEFAVVGSMPPSK--MHFGHKWVVDLLKFQKVTNINPI-----ADLEAYWA 112  
134 KDLTLQAYGDAV--ENAKDIIACGFDINKTFFISDLDMGSSGFYKWKI--QKHVTFN 191  
113 RNMSFTTKELALNEVITNYIALGLDPEKINVLQSKYQV-----KDLALILSKRTWS 167  
192 QVKGIFGFTDSCIGKISPPAIAQAPSFNSPFIQIPRDRIT--DIQCLIPCAIDQPYFRM 249  
168 EMKAIYGFKSGEINIGHVFAPIVQADIL---HPQLDENJSPKPVVVPVGVIGDQPHRL 224  
250 TRDVAPR---IGYKPKALLHSTFPFALQQAOTKMSADPNSSIFLTDITAKQIKTKVKA 306  
225 TRDIANRAKEFKTIPSSSTYHRWTGLGG--KWSKSPETALFLTDDEKTVKKIFS--A 281  
307 FSGGRDTIEHROFGG---NCDVNSVFWYLTFFLEDDDKLEIQINXQVTSGLMTGEXKA 363  
282 KTGGRTELEHKYGVGPEECWVVEFLY--HLIIDDKELAEIYQKCRSGELTGCCKKM 339  
364 LIEVLQPLIAHQARKEVTDIVK 388

340 AYERWVEFLKOLKEKREQAEXIAVK 364

RESULT 12  
B69131  
tryptophan-tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain Delta  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 03-Jun-2002  
C:Accession: B69131  
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F.  
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;  
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi  
A:Reference number: A69000; MUID:98037514; PMID:9317463  
A:Accession: B69131  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <MTH>  
A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84757.1; PID:g2621301  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH251  
A:Start codon: TTG  
C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 18.8%; Score 397.5; DB 2; Length 364;  
Best Local Similarity 27.9%; Pred. No. 1.1e-24;  
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;  
15 VDPWTVTSKAGIDYDKLIVRFSGSKIDKELINIERATGQPHFLRRGIFPFSHRDMN 74  
2 IDPW----GSAK--LEYQDLLENFGVRPF--SEVLDEV-----PEFSWLMRGLIIGHRDYE 50  
75 QVLDAVENKKPFYLYTGRGSPSEAMHVGHLIPFIETKWLQDVFNVLVIOQMTDDEKYLW 134  
51 RIISAMKKGEDFAVVTGMPSGR--MHIGHKMWVDQLW--YDRMGAEIPIPIADMEAYSAR 108  
135 DLTLQAYGDAVEN--AKDIIACGFDNK-----TFISDLDMGSSSGYKVKV 183  
109 GVDPEDSRR--AIIEYTAGYTAGLGLDEKDNHVVYLOSENLMLVEDLAYV----- 156  
184 IQKHVTFNWKVIGFTDSCIGKISPPAIAQAPSFNSPFIQIPRDRITDIQCLIPCAIDQ 243  
157 LAGKVNENELRALYGTGTS--SMAMYAPII--QVSDILHPQLDELGGPR---PVIVVGPDQ 213  
244 DPYFRMTDRVAPRI---GYPKALLHSTFPFALQQAOTKMSADPNSSIFLTDITAKQIK 299  
214 DPHIRLTRDIAARFRDRYGFILPSSTYHRFMGGLTGG--KWSNRPKSAIFLSDTPEAE 271  
300 TKYNKAFSGRDTIEHROFGNCDVDSFMYLTFFLE--DDKLEQIRKDYTSGLMTG 358  
272 AKI--RNAKTGRETLEKQRELGVPEECIIYETLLXHMSSGDSKLEELIYESCRNGTLMCG 330  
359 ELKKALIEVLQPLIAHQARKE 381  
331 ECKNNTAEFKKFEELSVREK 353

RESULT 13  
T43806  
tryptophan-tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)  
C:Species: Encephalitozoon cuniculi  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 03-Jun-2002  
C:Accession: T43806  
R:Peyretailade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.  
Mol. Biol. Evol. 15, 693-699, 1998  
A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene  
A:Reference number: 222693; MUID:98277683; PMID:9615449  
A:Accession: T43806  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PEY>

A:Cross-references: EMBL:AJ012470; PIDN:CAA10034.1

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

C:keywords: ligase

Query Match 18.2%; Score 386; DB 2; Length 134;  
Best Local Similarity 51.5%; Pred. No. 2.4e-24;  
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 12 EDFVDPWVOTSSAK---GIDYDKLIVFGSSKIDKELINRIERATGQRPHFLRGRIF 67  
DB 3 EORTPMDVEVVSDEVDVADYDKILNFCGEKNQALDRLEKLSGKPAHYFRGRIV 62  
QY 68 FSHRDNQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQPVFNPLVQMTD 127  
DB 63 FAHRDENLLDEIANNRRFYLTGRGSSKTHIGHTIPFLCLKYQDAFKRLVQITD 122  
QY 128 DEKYLKDLTLD 139  
DB 123 DEKFLWKSMLJE 134

RESULT 14

E:9461

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C:Accession: B69461

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

; Glodex, A.; Zhou, L.; Overbeek, J.D.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, R.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.

Saith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98C49343; PMID:9389475

A:Accession: B69461

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 <KLE>

A:Cross-references: GB:AE000782; NID:G2689309; PIDN:AA899554.1; PID:G264885

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 17.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 2.2e-22;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 15 VDPWVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHFLRGRIFFSHRDYN 74  
DB 3 VTPMEVEGV---IDYSKLIEFGMQPF-SBVLPEID-----NPHILMRGAIFGHRDYN 52  
QY 75 QVLDAVENKPKFYLTGRGSSSEAMHVGHLIPFTKWLQPVFNPLVQMTDDEKYLWK 134  
DB 53 RLIEAMQKPEPWAWSGFMPSG-LPHFGKHTMDLIVHQSAGKAFV-ALADMEAHSVR 110  
QY 135 DLTLDAQYGDVAENAKIICAGFDINKFTIFSD-LYMGSSGFGYKNV-KIQKHTVFNQV 193  
DB 111 GLSWEKTRGLMYIKSIILALGREDAVIYFQS-----KSHVKDLAFELSAEWNFSEL 164  
QY 194 KGIFGFTSDICIGKISFPAIQAPSFNSFPQIFDRDTIQCLIPCAIDQDPYRMTDV 253  
DB 165 RAIYGFNSDTSIAKMFVTAIQADIL---HPQLSDFGFGKPVVVPVGADQDPHMLTRDL 221  
QY 254 APRI----- 257  
DB 222 AARLISFSEFPEVGVRSRKGAELSSLDLEEDKKIYEHMDFGEAREIERAVRKI 281

QY 258 -----GYPKALLHSTFFPALQQAOTKMSASDPSNSIFLTDTAQIKTKVKNHAFSGG 310  
DB 282 EVRIGGFAPFIPSSYHRTFTGLTGG--KMSSKKEESYISLLDPPEGAKKVMK-AFTGG 338  
QY 311 RDTIEHRQFGNCVDVSYFMYLFFLED--DKLEQIRKDYTSGLMIGELKKALIEVLQ 369

DB 339 RATAEQRRLGCEPRCVVVELYSPHLIDSOBELNQIARCEGRLLCGCKKMAAELVK 398  
QY 370 PLIAHQARKEV 382  
DB 399 SFLKEHQERMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1;

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jit-no, K.; Takaha

wa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10362966

A:Accession: D72477

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KAM>

A:Cross-references: DBJ:AP000064; NID:G5105945; PIDN:BA81476.1; PID:G5106165

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2461

C:Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology

Query Match 16.7%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 5e-2;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 15 VDPWVOTSSAKGIDYDKLIVFGSSKIDKELINRIERATGQRPHFLRGRIFFSHRDYN 74  
DB 8 LDPW---GAVEIKDYKLLRTFGIRPSEVL--PLLKAGWEPSELKMRGIIFGHRDFD 61  
QY 75 QVLDAVENKPKFYLTGRGSSSEAMHVGHLIPFTKWLQ-DVFNVELVIQMTDDEKYLW 133  
DB 62 KILEAKARGERVAVLTGFMPGSK-FHFGKLTVDQLIYLQKNGKVV-FVAIADAAFAV 118  
QY 134 KDLTLDAQYGDVAEN-ACQIACGFDINKT-FIPSDLYMGSSGFGYKNVVIQKHVTEN 191  
DB 119 RRIIGEEAVRIAVEEYIANMIALGHDPKDTEFYQ---TNRGTYFRLIQLFSGKVTAA 174  
QY 192 OVKGIFG-FTSDICIGKISFPAIQAPSFNSFPQIFDRDTIQCLIPCAIDQDPYRMT 250  
DB 175 EMEATYGLTAKWASIT-----QAADILHVOLDVYGYR---HVVPVGADQDPHLRLT 227  
QY 251 RDVAER-----IGYKPAALLHSTFFPALQQAOTKMSASDPSNSIFLTDTAQIKTKVKNHA 306  
DB 228 RDLADRMAGVVELERFASTYHKLQPLDG--RKMSSSRPDSIFLTDPPPEVARNKLFR-A 284  
QY 307 PSGGRDITIEHRQFGNCV-DVSPMYLTFLEDDDKLEQIRKDYTS---GAMLTGELKK 362  
DB 285 ITGGRTAEEQRRLGGVPEVCVYHMDLYHMLPDDGVEVKHI---YTSRLGKILGCEQK 341

QY 363 ALIEVLQPLIAHQARRKRVTEIIVKESTPR 394

DB 342 IAWEKLERFLAHHQSRLEKAKTIANKVNEPPR 373

Search completed: August 24, 2004, 18:03:46

Job time : 43 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 24, 2004, 17:33:15 ; Search time 25 seconds  
(without alignments)  
835.205 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNRGPDATEAEEDFVDPWTV.....VTDEIVKFWTPKXLSFDFQ 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID        | Description        |
|------------|--------|-------------|--------|--------------|--------------------|
| 1          | 2101   | 99.3        | 471    | 1 SYW_HUMAN  | P23381 homo sapien |
| 2          | 2028.5 | 95.9        | 475    | 1 SYW_BOVIN  | P17248 bos taurus  |
| 3          | 1961.5 | 92.7        | 475    | 1 SYW_RABIT  | P23612 cryptolepus |
| 4          | 1948   | 91.6        | 481    | 1 SYW_MOUSE  | P2921 mus musculus |
| 5          | 1210   | 57.2        | 395    | 1 SYW_SCHPO  | Q09692 schizosacch |
| 6          | 1163   | 55.0        | 432    | 1 SYW_YEAST  | Q12109 saccharomyc |
| 7          | 910.5  | 43.0        | 381    | 1 SYW_SULTO  | Q976m1 sulfolobus  |
| 8          | 905    | 42.8        | 380    | 1 SYW_SULSO  | Q972x0 sulfolobus  |
| 9          | 826.5  | 39.1        | 385    | 1 SYW_PYRPU  | Q8u453 pyrococcus  |
| 10         | 803    | 37.9        | 385    | 1 SYW_PYRAB  | Q9uyl1 pyrococcus  |
| 11         | 796    | 37.6        | 386    | 1 SYW_PYRHO  | Q59584 pyrococcus  |
| 12         | 735    | 34.7        | 375    | 1 SYW_PYRAE  | Q8ztu5 pyrobaculum |
| 13         | 534.5  | 25.3        | 380    | 1 SYW2_HALN1 | Q9bn66 halobacteri |
| 14         | 453    | 21.4        | 374    | 1 SYW_METKA  | Q8tyf7 methanopyru |
| 15         | 409.5  | 19.4        | 370    | 1 SYW_METTA  | Q58810 methanococc |
| 16         | 397.5  | 18.8        | 364    | 1 SYW_METTH  | Q26352 methanobact |
| 17         | 386    | 18.2        | 134    | 1 SYW_ENCCU  | Q96771 encephalito |
| 18         | 370.5  | 17.5        | 420    | 1 SYW_ARCFU  | Q28579 archaeoglob |
| 19         | 353    | 16.7        | 374    | 1 SYW_AERPE  | Q9v924 aeropyrum p |
| 20         | 329.5  | 15.6        | 437    | 1 SYW_METAC  | Q8tual methanosarc |
| 21         | 313.5  | 14.8        | 491    | 1 SYW_METWA  | Q9pw55 methanosarc |
| 22         | 299.5  | 14.2        | 426    | 1 SYW_THEVO  | Q978x8 thermoplasm |
| 23         | 269.5  | 12.7        | 513    | 1 SYW1_HALN1 | Q9hn83 halobacteri |
| 24         | 268.5  | 12.7        | 426    | 1 SYW_THEAC  | Q9hiw5 thermoplasm |
| 25         | 203.5  | 9.6         | 341    | 1 SYW_STR3A  | Q8e2j5 streptococc |
| 26         | 203.5  | 9.6         | 341    | 1 SYW_STR3S  | Q8dwp7 streptococc |
| 27         | 192    | 9.1         | 323    | 1 SYW_ARCFU  | Q29482 archaeoglob |
| 28         | 185.5  | 8.8         | 341    | 1 SYW_STRPN  | Q97n42 streptococc |
| 29         | 183.5  | 8.7         | 340    | 1 SYW_STRPD  | Q8nyz2 streptococc |
| 30         | 182.5  | 8.6         | 340    | 1 SYW_STRPY  | Q99xb4 streptococc |
| 31         | 181    | 8.6         | 340    | 1 SYW_STRMU  | Q8dr1 streptococc  |
| 32         | 175.5  | 8.3         | 341    | 1 SYW_CLOIO  | Q46127 clostridium |
| 33         | 174.5  | 8.2         | 341    | 1 SYW_LACIA  | Q9c5d1 lactococcus |

## RESULT 1

### SYW\_HUMAN

ID SYW\_HUMAN STANDARD; PRT; 471 AA.  
AC P23381; P78535; Q9UDL3;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptostan--tRNA ligase)  
DE (TFPRS) (IFP53) (hWRS).  
GN WARS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OK NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92105071; PubMed=1761529;  
RA Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;  
RT "Interferon induces tryptophanyl-tRNA synthetase expression in human  
RT fibroblasts.";  
RL J. Biol. Chem. 266:24245-24248(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92107982; PubMed=1763065;  
RA Fleckner J., Rasmussen H.H., Justesen J.;  
RT "human interferon gamma potentially induces the synthesis of a 55-kDa  
RT protein [gamma 2] highly homologous to rabbit peptide chain release  
RT factor and bovine tryptophanyl-tRNA synthetase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92112058; PubMed=1765274;  
RA Frolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinzivleva O.L.,  
RA Kisselev L.L.;  
RT "Cloning and nucleotide sequence of the structural gene encoding for  
RT human tryptophanyl-tRNA synthetase.";  
RL Gene 109:291-296(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92164636; PubMed=1537332;  
RA Buwitt U., Flohr T., Boettger E.C.;  
RT "Molecular cloning and characterization of an interferon induced  
RT human cDNA with sequence homology to a mammalian peptide chain  
RT release factor.";  
RL EMBO J. 11:489-496(1992).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Xooe T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.Y., Hong L.,  
RA Stapleton M., Soares N.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carrinci P., Frange C.,

## ALIGNMENTS

|    |       |     |     |   |            |
|----|-------|-----|-----|---|------------|
| 34 | 172   | 8.1 | 394 | 1 | SYW_YEAST  |
| 35 | 170.5 | 8.1 | 366 | 1 | SYW_SULSO  |
| 36 | 169   | 8.0 | 395 | 1 | SYW_AQUAE  |
| 37 | 166.5 | 7.9 | 346 | 1 | SYW_CHLTR  |
| 38 | 163.5 | 7.7 | 353 | 1 | SYW_BORBU  |
| 39 | 159.5 | 7.5 | 337 | 1 | SYW_TREPA  |
| 40 | 158   | 7.5 | 346 | 1 | SYW_CHLMC  |
| 41 | 154   | 7.3 | 345 | 1 | SYW_CHLCV  |
| 42 | 152   | 7.2 | 344 | 1 | SYW_CHLPN  |
| 43 | 146   | 6.9 | 326 | 1 | SYW_HELPY  |
| 44 | 142.5 | 6.7 | 337 | 1 | SYW1_STR3M |
| 45 | 138   | 6.5 | 319 | 1 | SYW_METTH  |

P36421 saccharomyc  
P95982 sulfolobus  
O67115 aquifex aeo  
O84589 chl-amydia t  
O51038 borrelia bu  
O83640 treponema p  
Q82j15 chlamydia m  
Q821h9 chlamydia p  
Q927a4 chlamydia p  
P56396 helicobacte  
Q82hul streptomyc  
O27795 methanobact



Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahey S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." [6] *Proc. Natl. Acad. Sci. U.S.A.* 99:15899-15903 (2002).

SEQUENCE OF 1-13 FROM N.A. MEDLINE=96319994; PubMed=8724762; Sokolova I.V., Narovlianskii A.N., Achenkova A.M., Turpaev K.T.; "Alternative splicing of 5'-terminal exons of the human tryptophanyl-tRNA synthetase gene." [7] *Mol. Biol. (Mosk)* 30:319-329 (1996).

SEQUENCE OF 1-141 AND 182-471 FROM N.A. TISSUE-Sperm; MEDLINE=93292992; PubMed=7685728; Frolova L.Y., Grigorieva A.Y., Sudomoina M.A., Kisselev L.L.; "The human gene encoding tryptophanyl-tRNA synthetase: interferon-response elements and exon-intron organization." [8] *Gene* 128:237-245 (1993).

SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365: TISSUE=Keratinocytes; MEDLINE=93162043; PubMed=1286667; Rasmussen H.H., van Damme J., Fuype M., Gesser B., Celis J.E., Vandeckerckhove J.; "Microsequences of 145 proteins recorded in the two-dimensional gel protein database of normal human epidermal keratinocytes." [9] *Electrophoresis* 13:960-969 (1992).

FUNCTION. MEDLINE=92225128; PubMed=1373391; Bange F.-C., Flohr T., Buwitt U., Boettger E.C.; "An interferon-induced protein with release factor activity is a tryptophanyl-tRNA synthetase." [10] *FEBS Lett.* 300:162-166 (1992).

-!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP + diphosphate + L-tryptophanyl-tRNA(Trp).

-!- SUBUNIT: Homodimer.

-!- INDUCTION: By interferon gamma.

-!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

-!- SIMILARITY: Contains 1 WHEP-TRS domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M77804; AAA67324.1; -  
EMBL; X59892; CAA42545.1; -  
EMBL; M61715; AAA61298.1; -  
EMBL; X62570; CAA44450.1; -  
EMBL; BC017489; AAL17489.1; -  
EMBL; S82905; AAB39381.1; -  
EMBL; X67920; CAB94198.1; -  
EMBL; X67921; CAB94198.1; JOINED.  
EMBL; X67922; CAB94198.1; JOINED.  
EMBL; X67923; CAB94199.1; -  
EMBL; X67924; CAB94199.1; JOINED.  
EMBL; X67925; CAB94199.1; JOINED.  
EMBL; X67926; CAB94199.1; JOINED.  
EMBL; X67927; CAB94199.1; JOINED.

DR EMBL; X67928; CAB94199.1; JOINED.  
DR PIR; A41633; A41706.  
DR Aarhus/Ghent-2DPAGE; 3524; IEF.  
DR PHCI-2DPAGE; P23381; -.  
DR Genew; HGNC:12729; WARS.  
DR MIM; 191050; -.  
DR GO; GO:0005737; C-cytoplasm; TAS.  
DR GO; GO:0005625; C-soluble fraction; TAS.  
DR GO; GO:0008285; P-negative regulation of cell proliferation; TAS.  
DR GO; GO:0008412; P-protein biosynthesis; TAS.  
DR GO; GO:0008436; P-tryptophanyl-tRNA aminoacylation; TAS.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR002305; tRNA-synt\_1b.  
DR InterPro; IPR002306; tRNA-synt\_1b.  
DR InterPro; IPR002306; tRNA-synt\_1b.  
DR Pfam; PF00579; tRNA-synt\_1b; 1.  
DR Pfam; PF00458; WHEP-TRS; 1.  
DR PRINTS; PRO1039; TRNASYNTHTRP.  
DR TIGRFAMs; TIGR00233; tids; 1.  
DR PROSITE; PS02178; AA TRNA LIGASE\_1; 1.  
DR PROSITE; PS02762; WHEP-TRS; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
FT DOMAIN 19 64 WHEP-TRS.  
FT SITE 164 173 "HIGH" REGION.  
FT SITE 349 353 "KMSKS" REGION.  
FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
FT CONFLICT 424 424 A -> R (IN REF. 4).  
SQ SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;  
Query Match 99.3%; Score 2101; DB 1; Length 471;  
Best Local Similarity 99.5%; Pred. No. 2.3e-163;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
DB 71 SNHGPDATAEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
QY 61 FLRRGIFFSHRDMMQVLDAYENKPFYLYTGRPSSEAHVGHLPFFFTKWLQDVNPV 120  
DB 131 FLRRGIFFSHRDMMQVLDAYENKPFYLYTGRPSSEAHVGHLPFFFTKWLQDVNPV 190  
QY 121 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIACGFDINKTFFISDLDYMGSSGFYKN 180  
DB 191 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIACGFDINKTFFISDLDYMGSSGFYKN 250  
QY 181 VVKIQKHVTENQKVGIFGFTSDCGIKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 240  
DB 251 VVKIQKHVTENQKVGIFGFTSDCGIKISFPAIQAPSFNSFPQIFRDRDIOCLIPCA 310  
QY 241 IDQDPYFRMTDVAIPRIGYKPKALLHSTFFPALQGAQTMSASDNNSSIFLDTAKQIKT 300  
DB 311 IDQDPYFRMTDVAIPRIGYKPKALLHSTFFPALQGAQTMSASDNNSSIFLDTAKQIKT 370  
QY 301 KVNKHFSGGRTTIEHRQFGNCDDVSVFMYLFFLEDDKLEQIRKQYTGAMLTGEL 360  
DB 371 KVNKHFSGGRTTIEHRQFGNCDDVSVFMYLFFLEDDKLEQIRKQYTGAMLTGEL 430  
QY 361 KKALIEVLQPLIAEHQARKKVTDEIVKEFMTPRKLSDFQC 401  
DB 431 KKALIEVLQPLIAEHQARKKVTDEIVKEFMTPRKLSDFQC 471

## RESULT 2

SW\_BOVIN  
ID SW\_BOVIN STANDARD; ERT; 475 AA.  
AC P17248;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase) (TrpRS).  
GN WARS.  
OS Bos taurus (Bovine).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;

RN [1]  
 SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RP TISSUE=Retina;

RC MEDLINE=91329348; PubMed=1907847;

RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,

RA Gueguen M., Benedetto J.-P., Sallatranque M.-L., Alteriot J.,

RA Labouesse J., Bonnet J.,

RA Guenou M., Sarger C., Labouesse B., Bonnet J.,

RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to

RT prokaryotic synthetases but near identity with mammalian peptide

RT chain release factor.";

RL Biochemistry 30:7809-7817(1991).

RN [2]

SEQUENCE OF 17-475 FROM N.A.

RP TISSUE=Pancreas;

RC Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,

RA Gueguen M., Benedetto J.-P., Sarger C., Alteriot J., la Boussec B.,

RA Labouesse J., Bonnet J.,

RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

CC diphosphate + L-tryptophanyl-tRNA(Trp).

CC -!- SUBUNIT: Homodimer.

CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; X59318; CAA37872.1; -

EMBL; X52113; CAA36356.1; -

PIR; A40279; YWBO.

InterPro; IPR002305; tRNA-synt 1b.

InterPro; IPR001412; tRNA-synt 1.

InterPro; IPR002306; Trp tRNA-synt 1b.

InterPro; IPR000738; WHEP-TRS.

Pfam; PF00579; tRNA-synt 1b; 1.

Pfam; PF00458; WHEP-TRS; 1.

PRINTS; PR01039; TRNASYNTHTRP.

TIGRFAMs; TIGR00233; tips; 1.

PROSITE; PS00178; AA tRNA LIGASE\_I; 1.

PROSITE; PS02762; WHEP-TRS; 1.

K1 Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.

F1 DOMAIN 24 69 WHEP-TRS.

F1 SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.

F1 SITE 169 178 "HIGH" REGION.

F1 SITE 353 357 "KMSKS" REGION.

F1 CONFLICT 17 17 L -> M (IN REF. 2).

SEQUENCE 475 AA; S3729 MW; F7E531750137EB32 CRC64;

Query Match 95.9%; Score 2028.5; DB 1; Length 475;

Best Local Similarity 95.8%; Pred. No. 1.9e-157;

Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Q/ 1 SNHGDATEAEEDFVDPMTVQTSSAGKIDYKLVRFSSKIDKELINRIERATQGRPH 60

D/ 76 SGEGLDATEAEEDFVDPMTVQTSSAGKIDYKLVRFSSKIDKELINRIERATQGRPH 135

Q/ 61 FLRGIFPSHDMQVLDAENKPFYLYTGSPSSAMVGHLLPFIPTKWLQDVENV 120

D/ 136 FLRGIFPSHDMQVLDAENKPFYLYTGSPSSAMVGHLLPFIPTKWLQDVENV 195

Q/ 121 LVIQMTDEKYLWDLTDQAYGDAVENAKOIIACGFOINKTFFISDLDMGMSGFYKN 180

D/ 196 LVIQMTDEKYLWDLTDQAYGDAVENAKO-IICGFOINKTFFISDLDMGMSGFYKN 254

QY 181 VVKIQHVTNFKGIFQFTDSDCTGKISFPAIQAPSFNSFPQIFRDRDTQCLIPCA 240  
 DB 255 VVKIQHVTNFKGIFQFTDSDCTGKISFPAIQAPSFNSFPQIFRDRDTQCLIPCA 314  
 QY 241 IDODPYERWTRDVAIRIGYKPKALAHSTFFPALQGAOTKWSASDPSNIFLDTCAQIKT 300  
 DB 315 IDODPYERWTRDVAIRIGYKPKALAHSTFFPALQGAOTKWSASDPSNIFLDTCAQIKT 374  
 QY 301 KYNKEAFSGGRDTIEHRQFGNGCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 360  
 DB 375 KYNKEAFSGGRDTIEHRQFGNGCVDVSMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 434  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDIEIKFMTPRKLSYDFQ 401  
 DB 435 KKALIEVLQPLIAEHQARRKEVTDIEIKFMTPRKLSYDFQ 475

# RESULT 3

SWW\_RABIT

ID SYW\_RABIT STANDARD; PRT; 475 AA.

AC P23612; Q28607;

DT 01-NOV-1991 (Rel. 20, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)

DE (TRPS).

OS WARS.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI\_TaxID=9986;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=90239043; PubMed=2185472;

RA Lee C.C., Craig W.J., Muzny D.M., Harlow E., Caskey C.T.;

RT "Cloning and expression of a mammalian peptide chain release factor

RT with sequence similarity to tryptophanyl-tRNA synthetases.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).

RN [2]

REVISIONS TO 169-174 AND 227-228, AND FUNCTION.

RX MEDLINE=94009008; PubMed=8404867;

RA Prolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,

RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-J.;

RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA

RT synthetase are distinct proteins.";

RL EMBL J. 12:4013-4019(1993).

CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +

CC diphosphate + L-tryptophanyl-tRNA(Trp).

CC -!- SUBUNIT: Homodimer (by similarity).

CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.

CC -!- CAUTION: Was originally (Ref.1) thought to be a eukaryotic release

CC factor (BRP).

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; M33460; AAA31246.1; ALT SEQ.

EMBL; U02595; AAB60257.1; -

InterPro; IPR002305; tRNA-synt 1b.

InterPro; IPR001412; tRNA-synt 1.

InterPro; IPR002306; Trp tRNA-synt 1b.

InterPro; IPR000738; WHEP-TRS.

Pfam; PF00579; tRNA-synt 1b; 1.

Pfam; PF00458; WHEP-TRS; 1.

PRINTS; PR01039; TRNASYNTHTRP.

TIGRFAMs; TIGR00233; tips; 1.

PROSITE; PS00178; AA tRNA LIGASE\_I; 1.

```

DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;
Query Match 92.7%; Score 1961.5; DB 1; Length 475;
Best Local Similarity 91.5%; Pred. No. 5.3e-152;
Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;
QY 2 NHG-PDATTAEEDFVDPWTVTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60
DB 75 SHGDPEAVDDKEDFVDPWTVRTSSAKGIDYDKLIVRFGSSKIDKELVNIERATGQRPHR 134
QY 61 FRRGGTFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMHVGHLLIPFTKWLQDVFNVP 120
DB 135 FRRGGTFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMHVGHLLIPFTKWLQDVFNVP 194
QY 121 LVQMTDDKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFFSDLDYMGSSGFYKN 180
DB 195 LVVQMSDDKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFFSDLDYMGSSGFYKN 254
QY 181 VVKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCA 240
DB 255 VVKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCA 314
QY 241 IQDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKT 300
DB 315 IQDDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKT 374
QY 301 KYNKHAFGSGRDTIEHRQFGNCDVDVDFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 360
DB 375 KYNKHAFGSGRDTIEHRQFGNCDVDVDFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 434
QY 361 KALIEVLOPLIAEHQARRKEVTDEIVKFPMTPKLSFDFQ 401
DB 435 KXELIDVLOPEVAEHQARRKEVTDEIVKFPMTPKLSFDFQ 475

```

## RESULT 4

```

SYN_MOUSE STANDARD; PRT; 481 AA.
P32921;
01-OCT-1993 (Rel. 27, Created)
01-OCT-1993 (Rel. 27, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
(TIPS).
WARS OR WRS.
Mus musculus (Mouse).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
MEDLINE=95018226; PubMed=7932716;
Pajot B., Sarger C., Bonnet J., Garret M.;
"An alternative splicing modifies the C-terminal end of tryptophanyl-
tRNA synthetase in murine embryonic stem cells.";
J. Mol. Biol. 242:599-603(1994).
-!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
-!- SUBUNIT: Homodimer (By similarity).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Long;
CC IsoId=P32921-1; Sequence=Displayed;
CC Name=2; Synonyms=Short;
CC IsoId=P32921-2; Sequence=VSP_006313;
-!- TISSUE SPECIFICITY: Isoform 2 is widely expressed, isoform 1 is
found only in embryonic stem cells.
-!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.

```

```

-!- SIMILARITY: Contains 1 WHEP-TRS domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; X69656; CAA49347.1; -
EMBL; X69657; CAA49348.1; -
PIR; S50053; S50053.
MGD; MGI:104630; Wars.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001432; tRNA-synt_1.
DR InterPro; IPR002306; Trp-tRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRfam; TIGR00233; tRps; 1.
DR PROSITE; PS00178; AA-tRNA_LIGASE_I; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPLIC 476 481 Missing (in isoform 2).
FT /FTID=VSP_006313.
SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;
Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 4.4e-150;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;
QY 2 NHGPDATAEEDFVDPWTVTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61
DB 76 NCSDATKASEDFVDPWTVTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 135
QY 62 LRRGIFPSHRDMNQVLDAYENKPPVLYTGRGSSSEAMHVGHLLIPFTKWLQDVFNVP 121
DB 136 LRRGIFPSHRDMNQVLDAYENKPPVLYTGRGSSSEAMHVGHLLIPFTKWLQDVFNVP 195
QY 122 VIQMTDDKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFFSDLDYMGSSGFYKN 181
DB 196 VIQMTDDKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFFSDLDYMGSSGFYKN 255
QY 182 VKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCAI 241
DB 256 VKIQHVTFNQKGFSGTSDSCIGKISFPALCAAPSFNSPQIPRDRDTQCLIPCAI 315
QY 242 DQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKTK 301
DB 316 DQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSADPNSSIFLDTAKQIKTK 375
QY 302 VNHKAFSGRDTIEHRQFGNCDVDVDFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 361
DB 376 VNHKAFSGRDTIEHRQFGNCDVDVDFMYLTFLEDDDKLEQIRKDYTGAMLTGEL 435
QY 362 KALIEVLOPLIAEHQARRKEVTDEIVKFPMTPKLSFDFQ 401
DB 436 KTLIDVLOPLIAEHQARRKEVTDEIVKFPMTPKLSFDFQ 475

```

## RESULT 5

```

SYN_MOUSE STANDARD; PRT; 395 AA.
ID SYW SCHPO
AC C09592;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA  
ligase) (TrpRS).

GN SPAC2F7.13C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Williams R., Rajendram M.A., Lyne M., Stewart A.,  
RA Sgourou J., Peat M., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Fraser A.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy J., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler R., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Fersburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
FJ "The genome sequence of Schizosaccharomyces pombe.";  
FJ Nature 415:871-880(2002).  
C1 -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
C2 diphosphate + L-tryptophanyl-tRNA(Trp).  
C3 -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
C4  
C5  
C6  
C7  
C8  
C9  
C10  
C11  
C12  
C13  
C14  
C15  
C16  
C17  
C18  
C19  
C20  
C21  
C22  
C23  
C24  
C25  
C26  
C27  
C28  
C29  
C30  
C31  
C32  
C33  
C34  
C35  
C36  
C37  
C38  
C39  
C40  
C41  
C42  
C43  
C44  
C45  
C46  
C47  
C48  
C49  
C50  
C51  
C52  
C53  
C54  
C55  
C56  
C57  
C58  
C59  
C60  
C61  
C62  
C63  
C64  
C65  
C66  
C67  
C68  
C69  
C70  
C71  
C72  
C73  
C74  
C75  
C76  
C77  
C78  
C79  
C80  
C81  
C82  
C83  
C84  
C85  
C86  
C87  
C88  
C89  
C90  
C91  
C92  
C93  
C94  
C95  
C96  
C97  
C98  
C99  
C100  
C101  
C102  
C103  
C104  
C105  
C106  
C107  
C108  
C109  
C110  
C111  
C112  
C113  
C114  
C115  
C116  
C117  
C118  
C119  
C120  
C121  
C122  
C123  
C124  
C125  
C126  
C127  
C128  
C129  
C130  
C131  
C132  
C133  
C134  
C135  
C136  
C137  
C138  
C139  
C140  
C141  
C142  
C143  
C144  
C145  
C146  
C147  
C148  
C149  
C150  
C151  
C152  
C153  
C154  
C155  
C156  
C157  
C158  
C159  
C160  
C161  
C162  
C163  
C164  
C165  
C166  
C167  
C168  
C169  
C170  
C171  
C172  
C173  
C174  
C175  
C176  
C177  
C178  
C179  
C180  
C181  
C182  
C183  
C184  
C185  
C186  
C187  
C188  
C189  
C190  
C191  
C192  
C193  
C194  
C195  
C196  
C197  
C198  
C199  
C200  
C201  
C202  
C203  
C204  
C205  
C206  
C207  
C208  
C209  
C210  
C211  
C212  
C213  
C214  
C215  
C216  
C217  
C218  
C219  
C220  
C221  
C222  
C223  
C224  
C225  
C226  
C227  
C228  
C229  
C230  
C231  
C232  
C233  
C234  
C235  
C236  
C237  
C238  
C239  
C240  
C241  
C242  
C243  
C244  
C245  
C246  
C247  
C248  
C249  
C250  
C251  
C252  
C253  
C254  
C255  
C256  
C257  
C258  
C259  
C260  
C261  
C262  
C263  
C264  
C265  
C266  
C267  
C268  
C269  
C270  
C271  
C272  
C273  
C274  
C275  
C276  
C277  
C278  
C279  
C280  
C281  
C282  
C283  
C284  
C285  
C286  
C287  
C288  
C289  
C290  
C291  
C292  
C293  
C294  
C295  
C296  
C297  
C298  
C299  
C300  
C301  
C302  
C303  
C304  
C305  
C306  
C307  
C308  
C309  
C310  
C311  
C312  
C313  
C314  
C315  
C316  
C317  
C318  
C319  
C320  
C321  
C322  
C323  
C324  
C325  
C326  
C327  
C328  
C329  
C330  
C331  
C332  
C333  
C334  
C335  
C336  
C337  
C338  
C339  
C340  
C341  
C342  
C343  
C344  
C345  
C346  
C347  
C348  
C349  
C350  
C351  
C352  
C353  
C354  
C355  
C356  
C357  
C358  
C359  
C360  
C361  
C362  
C363  
C364  
C365  
C366  
C367  
C368  
C369  
C370  
C371  
C372  
C373  
C374  
C375  
C376  
C377  
C378  
C379  
C380  
C381  
C382  
C383  
C384  
C385  
C386  
C387  
C388  
C389  
C390  
C391  
C392  
C393  
C394  
C395  
C396  
C397  
C398  
C399  
C400  
C401  
C402  
C403  
C404  
C405  
C406  
C407  
C408  
C409  
C410  
C411  
C412  
C413  
C414  
C415  
C416  
C417  
C418  
C419  
C420  
C421  
C422  
C423  
C424  
C425  
C426  
C427  
C428  
C429  
C430  
C431  
C432  
C433  
C434  
C435  
C436  
C437  
C438  
C439  
C440  
C441  
C442  
C443  
C444  
C445  
C446  
C447  
C448  
C449  
C450  
C451  
C452  
C453  
C454  
C455  
C456  
C457  
C458  
C459  
C460  
C461  
C462  
C463  
C464  
C465  
C466  
C467  
C468  
C469  
C470  
C471  
C472  
C473  
C474  
C475  
C476  
C477  
C478  
C479  
C480  
C481  
C482  
C483  
C484  
C485  
C486  
C487  
C488  
C489  
C490  
C491  
C492  
C493  
C494  
C495  
C496  
C497  
C498  
C499  
C500  
C501  
C502  
C503  
C504  
C505  
C506  
C507  
C508  
C509  
C510  
C511  
C512  
C513  
C514  
C515  
C516  
C517  
C518  
C519  
C520  
C521  
C522  
C523  
C524  
C525  
C526  
C527  
C528  
C529  
C530  
C531  
C532  
C533  
C534  
C535  
C536  
C537  
C538  
C539  
C540  
C541  
C542  
C543  
C544  
C545  
C546  
C547  
C548  
C549  
C550  
C551  
C552  
C553  
C554  
C555  
C556  
C557  
C558  
C559  
C560  
C561  
C562  
C563  
C564  
C565  
C566  
C567  
C568  
C569  
C570  
C571  
C572  
C573  
C574  
C575  
C576  
C577  
C578  
C579  
C580  
C581  
C582  
C583  
C584  
C585  
C586  
C587  
C588  
C589  
C590  
C591  
C592  
C593  
C594  
C595  
C596  
C597  
C598  
C599  
C600  
C601  
C602  
C603  
C604  
C605  
C606  
C607  
C608  
C609  
C610  
C611  
C612  
C613  
C614  
C615  
C616  
C617  
C618  
C619  
C620  
C621  
C622  
C623  
C624  
C625  
C626  
C627  
C628  
C629  
C630  
C631  
C632  
C633  
C634  
C635  
C636  
C637  
C638  
C639  
C640  
C641  
C642  
C643  
C644  
C645  
C646  
C647  
C648  
C649  
C650  
C651  
C652  
C653  
C654  
C655  
C656  
C657  
C658  
C659  
C660  
C661  
C662  
C663  
C664  
C665  
C666  
C667  
C668  
C669  
C670  
C671  
C672  
C673  
C674  
C675  
C676  
C677  
C678  
C679  
C680  
C681  
C682  
C683  
C684  
C685  
C686  
C687  
C688  
C689  
C690  
C691  
C692  
C693  
C694  
C695  
C696  
C697  
C698  
C699  
C700  
C701  
C702  
C703  
C704  
C705  
C706  
C707  
C708  
C709  
C710  
C711  
C712  
C713  
C714  
C715  
C716  
C717  
C718  
C719  
C720  
C721  
C722  
C723  
C724  
C725  
C726  
C727  
C728  
C729  
C730  
C731  
C732  
C733  
C734  
C735  
C736  
C737  
C738  
C739  
C740  
C741  
C742  
C743  
C744  
C745  
C746  
C747  
C748  
C749  
C750  
C751  
C752  
C753  
C754  
C755  
C756  
C757  
C758  
C759  
C760  
C761  
C762  
C763  
C764  
C765  
C766  
C767  
C768  
C769  
C770  
C771  
C772  
C773  
C774  
C775  
C776  
C777  
C778  
C779  
C780  
C781  
C782  
C783  
C784  
C785  
C786  
C787  
C788  
C789  
C790  
C791  
C792  
C793  
C794  
C795  
C796  
C797  
C798  
C799  
C800  
C801  
C802  
C803  
C804  
C805  
C806  
C807  
C808  
C809  
C810  
C811  
C812  
C813  
C814  
C815  
C816  
C817  
C818  
C819  
C820  
C821  
C822  
C823  
C824  
C825  
C826  
C827  
C828  
C829  
C830  
C831  
C832  
C833  
C834  
C835  
C836  
C837  
C838  
C839  
C840  
C841  
C842  
C843  
C844  
C845  
C846  
C847  
C848  
C849  
C850  
C851  
C852  
C853  
C854  
C855  
C856  
C857  
C858  
C859  
C860  
C861  
C862  
C863  
C864  
C865  
C866  
C867  
C868  
C869  
C870  
C871  
C872  
C873  
C874  
C875  
C876  
C877  
C878  
C879  
C880  
C881  
C882  
C883  
C884  
C885  
C886  
C887  
C888  
C889  
C890  
C891  
C892  
C893  
C894  
C895  
C896  
C897  
C898  
C899  
C900  
C901  
C902  
C903  
C904  
C905  
C906  
C907  
C908  
C909  
C910  
C911  
C912  
C913  
C914  
C915  
C916  
C917  
C918  
C919  
C920  
C921  
C922  
C923  
C924  
C925  
C926  
C927  
C928  
C929  
C930  
C931  
C932  
C933  
C934  
C935  
C936  
C937  
C938  
C939  
C940  
C941  
C942  
C943  
C944  
C945  
C946  
C947  
C948  
C949  
C950  
C951  
C952  
C953  
C954  
C955  
C956  
C957  
C958  
C959  
C960  
C961  
C962  
C963  
C964  
C965  
C966  
C967  
C968  
C969  
C970  
C971  
C972  
C973  
C974  
C975  
C976  
C977  
C978  
C979  
C980  
C981  
C982  
C983  
C984  
C985  
C986  
C987  
C988  
C989  
C990  
C991  
C992  
C993  
C994  
C995  
C996  
C997  
C998  
C999  
C1000

EMBL; Z50142; CAA90500.1; --  
PIR; S58157; S58157.  
GeneDB SPombe; SPAC2F7.13C; --  
InterPro; IPR002305; tRNA-synt\_1b.  
InterPro; IPR001412; tRNA-synt\_1.  
InterPro; IPR002306; Trp tRNA-synt\_1b.  
Pfam; PF00579; tRNA-synt\_1b; 1.  
PRINTS; PR01039; TRNASYNTHTRP.  
TIGRFAMs; TIGR00233; ttps; 1.  
PROSITE; PS00178; AA TRNA LIGASE I; 1.  
KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;  
Ligase; ATP-binding.  
FT SITE 91 100 "HIGH" REGION.  
FT SITE 275 279 "RMSKS" REGION.  
SQ SEQUENCE 395 AA; 44910 MW; E656A8B76C5FDF9 CRC64;  
Query Match 57.2%; Score 1210; DB 1; Length 395;  
Best Local Similarity 59.6%; Pred. No. 5.7e-91;  
Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

Qy 11 BEDFVDPWTQTS-----SAKGDYDKLIVRFSSKIDKELINIERATQCPHHFLRGI 66  
Db 4 BEQIVTPDWKGSIVDGEKGDYELVQFGRKTIPTQLERFEKLGKPKHLLIRGA 63

QY 67 FESHRDNQVLDAYENKKPFYLYTGRPSSSEAMVGHLLIPFTKWLQDVENVELVQMT 126  
Db 64 FESHRDFMLIDRYEQKKPFYLYTGRPSSSDSMHGHMIPFMCKWLQDVFEVQLVQLT 123  
QY 127 DDEKYLWKD-LTLDQAYGDVAENAKDIIACGFDINKTFIESDLDYMGSSGFYKNVYKIQ 185  
Db 124 DDEKFLKQGVSLDCCORFARENAXDIIIVGDFEKKTFIFMNSTYVG--GAFYQNVVRIA 181  
QY 186 KHTFNVQKGIPOFTSDSCIGKISFPALQAPZFSNSFPQIFRDRTPICQLIPCAIDQDP 245  
Db 182 KCITANGKACFGFTSDSISGKIHFASIQAPZFSNSFPQIFRDRTPICQLIPCAIDQDP 241  
QY 246 YFMTVDVABRIGYPKPALHSTFFPALQAGQCKMSASDPSNSTFLDTAKQIKYKXKH 305  
Db 242 YFRLTRDVSRLPKPKPALHSTFFPALQAGQCKMSASDPSNSTFLDTAKQIKYKXKH 301  
QY 306 AFSGGSDTIEHHRQFGNGCDVDSFMYLTFFLEDDKLEQIRKDYTGAMLTGELKKALI 365  
Db 302 AFSGGGATIEHREKGNPDVDVAYQVLSFFLDDDEKLEKLYNTYKAGTSLTGEMKGECI 361  
QY 366 EVLQPLIAHQARKEVTEIVKFMTP-PRKLSF 398  
Db 362 KLLQQFVSDFQAARSKYDEATLDMFMDGSRKLEW 395

## RESULT 6

SYNC\_YEAST STANDARD; PRE; 432 AA.  
ID SYNC\_YEAST  
AC 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--  
tRNA ligase) [TrpRS].  
GN WRS1 OR YOL097C OR HRE432.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96076631; PubMed=7502582;  
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;  
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV  
including the Trp-H3 retrotransposon, the sufi(+) frameshift  
suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a  
delta element.";  
RT Yeast 11:1069-1075(1995).  
RL Yeast [2]  
RF FUNCTION  
RE MEDLINE=97197969; PubMed=9046085;  
RA John T.R., Ghosh M., Johnson J.D.;  
RT "Identification and expression of the Saccharomyces cerevisiae  
cytoplasmic tryptophanyl-tRNA synthetase gene.";  
RL Yeast 13:37-41(1997).  
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z48149; CAA98164.1; --  
CC EMBL; Z74839; CAA99110.1; --  
CC PIR; S51901; S51901.

"Complete genome sequence of an aerobic thermoacidophilic

```

OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
R3 [1]
R2 SEQUENCE FROM N.A.
R1 STRAIN=ATCC 35092 / DSM 1617 / P2;
R0 MEDLINE=21332296; PubMed=11427726;
R3 She Q., Singh R.K., Confalocieri F., Zivanovic Y., Allard G.,
R2 Aways M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
R1 De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
R0 Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
R3 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
R2 Charlebois R.L., Doolittle W.F., Duguet X., Gaasterland T.,
R1 Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
R0 "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
R3 Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840 (2001).
R2 -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
R1 diphosphate + L-tryptophanyl-tRNA(Trp).
R0 -!- SUBCELLULAR LOCATION: Cytoplasmic.
C3 -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
C2 -----
C1 This SWISS-PROT entry is copyright. It is produced through a collaboration
C0 between the Swiss Institute of Bioinformatics and the EMBL outstation -
C3 the European Bioinformatics Institute. There are no restrictions on its
C2 use by non-profit institutions as long as its content is in no way
C1 modified and this statement is not removed. Usage by and for commercial
C0 entities requires a license agreement (See http://www.isb-sib.ch/announce/
C3 or send an email to license@isb-sib.ch).
C2 -----
DR EMBL; AE006677; AAK40778.1; AUT_INIT.
DR HAMAP; MF_00140; -.
DR InterPro; IPR002305; tRNA-synt 1b.
DR InterPro; IPR001412; tRNA-synt 1.
DR Pfam; PF00579; tRNA-synt 1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 81 89 "HIGH" REGION.
FT SITE 253 257 "KMSKS" REGION.
SQ SEQUENCE 360 AA; 44691 MW; CF8344CF63883580 CRC64;
Query Match 42.8%; Score 905; DB 1; Length 380;
Best Local Similarity 48.7%; Pred. No. 3.5e-66;
Matches 190; Conservative 68; Mismatches 108; Indels 24; Gaps 10;
QY 11 BEDFVDPWTVTSSAKG-IDYDKLIVRGSSKIDKELINRIERATGQPHHFRGIFPS 69
Db 3 DEFTVTPEV-----KGKVDYDKLIVQFGTQKITEELKQRIKLAGDL-HVMLRENVEFS 56
QY 70 HRDMNQVLDAYENKPPFLYLTGRGPSSSEAMHVGHLIPFTKWLQDVNPVLVITQDDE 129
Db 57 HRDLVLNDYKSKGFLYLTGRAPSL-GMIGHILIPFTKWLQKFNANLYEITDDE 115
QY 130 KYLWK-DLTLQAYGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKNVVKIQKHV 188
Db 116 KYMENPEFTLDQTSWAVNDILDIIVGDFEKTFIFQDTEYI---RNWYPTVYIAKL 172
QY 189 TFNQVKGIFGTFDSDCIKISFPALQAPSNPSPPQIFRORTDIQCLIPCAIDQDPYR 248
Db 173 TFSVRATFGLDASSNIGLIFYPALQIAPT-----MFKK--RCLIPAGIDQDPYR 222
QY 249 MTRDVAIPRIGPKALLHSTFFPALQCGQTQMSASDPNSSIFLDTAKIQTKNKHAFS 308
Db 223 LQDIALSLGYKAAQIHSKFLPPTGPEGKSSSNPETALYVDVDPKTVKRMKYAFS 282
QY 309 GGRDTIBEHQFGGNCVDVDFWMLTFFLEDD-KLEQIRKDYTSAGMLTGELKALLIEV 367
Db 283 GGOPTIELHRKYGNPEIDVDFQWLYYFEEDNNRIKIEBEYRSGKMLTGELKQLIDK 342

```

```

QY 368 LQPLIAEHQARKEKVEIDIVKEFMTPRKLS 397
Db 343 LNNFLEHR-RRREAEKELVHVFKYDGKLA 371

```

## RESULT 9

```

SYN_PYRFU
ID AC Q8U453; STANDARD; PRT; 385 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TPRS).
GN TRPS OR PFC241.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ Databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
C3 This SWISS-PROT entry is copyright. It is produced through a collaboration
C2 between the Swiss Institute of Bioinformatics and the EMBL outstation -
C1 the European Bioinformatics Institute. There are no restrictions on its
C0 use by non-profit institutions as long as its content is in no way
C3 modified and this statement is not removed. Usage by and for commercial
C2 entities requires a license agreement (See http://www.isb-sib.ch/announce/
C1 or send an email to license@isb-sib.ch).
C0 -----
DR EMBL; AE010149; AAL80365.1; -.
DR HAMAP; MF_00140; -.
DR InterPro; IPR002305; tRNA-synt 1b.
DR InterPro; IPR001412; tRNA-synt 1.
DR Pfam; PF00579; tRNA-synt 1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trps; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 82 90 "HIGH" REGION.
FT SITE 253 257 "KMSKS" REGION.
SQ SEQUENCE 385 AA; 45178 MW; 3A7A628958200CCC CRC64;

```

```

Query Match 39.1%; Score 826.5; DB 1; Length 385;
Best Local Similarity 45.9%; Pred. No. 8.8e-60;
Matches 180; Conservative 69; Mismatches 116; Indels 27; Gaps 10;

```

```

QY 11 BEDFVDPWTVTSSAKG-IDYDKLIVRGSSKIDKELINRIERATGQPHHFRGIFPS 69
Db 2 BEEFKVTPEVEGI-----IDYKLTIEQFGTSPDLDLERTARLTAKSELPIFRKRFPS 57
QY 70 HRDMNQVLDAYENKPPFLYLTGRGPSSSEAMHVGHLIPFTKWLQDVNPVLVITQDDE 129
Db 58 HRDYDKVLDDVEQKGFYLTGRGFSG-PMHIGHLIPFPAIKWLQKEDPNLYIQTDD 116
QY 130 KYLWK-DLTLQAYGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKNVVKIQKHV 188
Db 117 KFLFENLTFTEDTKYAVQCNILDIIVGDFEKTFIFQNSSEF---TKIYEMAIPTAKKI 172
QY 189 TFNQVKGIFGTFDSDCIKISFPALQAPSNPSPPQIFRORTDIQCLIPCAIDQDPYR 248
Db 173 NFSMAKAVFGFTEQSKIGNIFFPALQIAPT-----FEKK--RCLIPAIIDQDPYR 222

```

QY 249 MTRDVAIRIGYKPKALLHSTFFPALQAGTQMSASDPNSS:FLTDTAKQIKTKVKNKHAFS 308  
 DT 223 LQDPAESLGYKYTAALHSKFFPLTLGLEKMSASKPETA:YLTNDPEEACKLIWKFAIT 282  
 QY 309 GGRDITTEHRQFGNCVDVSEWYLTFEEDDDKLEQIRKDY---TSGAMLTGELKKALI 365  
 DT 283 GGGPTLKQREKGNPKCVVFKWLEIFFEDDK--KLMEYYACKNGELGCECKRYLI 340  
 QY 366 EVLQPLIAHQARKEVTDIVKEFMPTRKLS 397  
 DT 341 QRVOEFLKEHQRKK-AEKLVEKFKYTGKLA 371

RESULT 10  
 SYM PYRAB STANDARD; PRT; 385 AA.  
 AC Q9UY11;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TPRS)  
 DE TRPS OR PYRAB16970 OR PAB1111.  
 OS Pyrococcus abyssi.  
 OS Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OC NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GE5 / Orsay;  
 RX MEDLINE=22511545; PubMed=12622808;  
 RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,  
 RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,  
 RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;  
 RA "An integrated analysis of the genome of the hyperthermophilic  
 RA archaeon Pyrococcus abyssi.";  
 RT Mol. Microbiol. 47:1495-1512(2003).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AJ248288; CAB50601.1; -;  
 CC F01; C75020; C75020.  
 CC HAMAP; MF 00140; -; 1.  
 CC InterPro; IPR002305; tRNA-synt\_1b.  
 CC InterPro; IPR001412; tRNA-synt\_1.  
 CC Pfam; PF00579; tRNA-synt\_1b; 1.  
 CC PRINTS; PR01039; TRNASYNTHTRP.  
 CC TIGRFAMs; TIGR00233; tps; 1.  
 CC PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 CC Complete proteome.  
 CC SITE 82 30 "HIGH" REGION.  
 CC SITE 253 257 "RMSKS" REGION.  
 CC SEQUENCE 385 AA; 45100 MW; 4C29D01414976B12 CRC64;  
 Query Match 37.9%; Score 803; DB 1; Length 385;  
 Best Local Similarity 45.3%; Pred. No. 7.2e-58;  
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;  
 12 EDF-VDPWTQTSSAGKIDYDKLIVRFSSKIDKELINRIERATQRPHPHURRGIFFSH 70  
 3 EDFKVTPEWESGV---VDYKLIIEHFGTSPLTESLLEKTAELTKSELPLFFPRKFFFSH 58

QY 71 RDMQVLDAYENKKPFVLYTGRGSPSEMHVGHILPFIETKWLQOVENVPLVIQMTDDEK 130  
 DB 59 RDYDKVLDQTEEGRGFFLYTGRGSG-PMHIGHLIPFFATKWLQKFGVNLVIQITDDEK 117  
 QY 131 YLWKD-LTLQAYGDAVENAKDI-ACGFDINKTIFSDLDYMGSSGFYKNWVKIQKHVT 189  
 DB 118 FLFKNLTFEDTKHWAENILDIIVAGFDPDKTIFQNSEF---TKIYENAIPIAKKIN 173  
 QY 190 FNOVKHFGTSDSDICGISFPAIQAAAPSFNSPQIFRORTDIQCLIPCAIDQDPYPRM 249  
 DB 174 FSNKAVFGFTQSKIGMIFPAIQIAPTF-----FEKR---RLIIPAIADQDPYWR 223  
 QY 250 TRDVAPRIGYKPKALLHSTFFPALQAGTQMSASDPNSS:FLTDTAKQIKTKVKNKHAFS 309  
 DB 224 QRDFAESLGYKYTAALHSKFPVSLTSLGMSASKPETA:YLTNDPEEACKLIWKFAIT 283  
 QY 310 GRDTIEHRQFGNCVDVSEWYLTFEEDDDKLEQIRKDY---TSGAMLTGELKKALIE 366  
 DB 284 GRPTLKEQREKGEPEKCVVFKWLEIFFEEDDK--KLMEYYACKNGELGCECKRYLI 341  
 QY 367 VLQPLIAHQARKEVTDIVK 388  
 DB 342 KIQEFLKEHQRRKKAQKQIEK 363

RESULT 11  
 SYM PYRHO STANDARD; PRT; 386 AA.  
 AC O59584;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TPRS)  
 DE TRPS OR PH1921.  
 OS Pyrococcus horikoshii.  
 OS Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OC NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Negai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa E., Takamiya M., Ohfuka Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AP000007; BAA31046.1; ALT INIT.  
 CC HAMAP; MF 00140; -; 1.  
 CC InterPro; IPR002305; tRNA-synt\_1b.  
 CC InterPro; IPR001412; tRNA-synt\_1.  
 CC InterPro; IPR002306; Trp tRNA-synt\_1b.  
 CC Pfam; PF00579; tRNA-synt\_1b; 1.  
 CC PRINTS; PR01039; TRNASYNTHTRP.







Db 273 AKTGGALEBEREGGNFDECVVYELMHVLAIRIGGDEKLREIRKKCREGDIIOGECK 332  
Qy 362 KALIEVLQPLIAEHOARKEVTDEI 386  
Db 333 RMVGEALAEILEELERREDVRDEL 357

RESULT 15  
SYN\_METUA  
ID SYW METJA STANDARD; PRT; 370 AA.  
AC Q58810;  
DI 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TPRS).  
GN TRPS OR M11415.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-P., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
R1 \*Complete genome sequence of the methanogenic archaeon, Methanococcus  
R2 jannaschii";  
R3 Science 273:1058-1073(1996).  
R4  
C1 -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
C2 diphosphate + L-tryptophanyl-tRNA(Trp).  
C3 -I- SUBCELLULAR LOCATION: Cytoplasmic.  
C4  
C5 -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.  
C6  
C7  
C8  
C9  
C10  
C11  
C12  
C13  
C14  
C15  
C16  
C17  
C18  
C19  
C20  
C21  
C22  
C23  
C24  
C25  
C26  
C27  
C28  
C29  
C30  
C31  
C32  
C33  
C34  
C35  
C36  
C37  
C38  
C39  
C40  
C41  
C42  
C43  
C44  
C45  
C46  
C47  
C48  
C49  
C50  
C51  
C52  
C53  
C54  
C55  
C56  
C57  
C58  
C59  
C60  
C61  
C62  
C63  
C64  
C65  
C66  
C67  
C68  
C69  
C70  
C71  
C72  
C73  
C74  
C75  
C76  
C77  
C78  
C79  
C80  
C81  
C82  
C83  
C84  
C85  
C86  
C87  
C88  
C89  
C90  
C91  
C92  
C93  
C94  
C95  
C96  
C97  
C98  
C99  
C100  
C101  
C102  
C103  
C104  
C105  
C106  
C107  
C108  
C109  
C110  
C111  
C112  
C113  
C114  
C115  
C116  
C117  
C118  
C119  
C120  
C121  
C122  
C123  
C124  
C125  
C126  
C127  
C128  
C129  
C130  
C131  
C132  
C133  
C134  
C135  
C136  
C137  
C138  
C139  
C140  
C141  
C142  
C143  
C144  
C145  
C146  
C147  
C148  
C149  
C150  
C151  
C152  
C153  
C154  
C155  
C156  
C157  
C158  
C159  
C160  
C161  
C162  
C163  
C164  
C165  
C166  
C167  
C168  
C169  
C170  
C171  
C172  
C173  
C174  
C175  
C176  
C177  
C178  
C179  
C180  
C181  
C182  
C183  
C184  
C185  
C186  
C187  
C188  
C189  
C190  
C191  
C192  
C193  
C194  
C195  
C196  
C197  
C198  
C199  
C200  
C201  
C202  
C203  
C204  
C205  
C206  
C207  
C208  
C209  
C210  
C211  
C212  
C213  
C214  
C215  
C216  
C217  
C218  
C219  
C220  
C221  
C222  
C223  
C224  
C225  
C226  
C227  
C228  
C229  
C230  
C231  
C232  
C233  
C234  
C235  
C236  
C237  
C238  
C239  
C240  
C241  
C242  
C243  
C244  
C245  
C246  
C247  
C248  
C249  
C250  
C251  
C252  
C253  
C254  
C255  
C256  
C257  
C258  
C259  
C260  
C261  
C262  
C263  
C264  
C265  
C266  
C267  
C268  
C269  
C270  
C271  
C272  
C273  
C274  
C275  
C276  
C277  
C278  
C279  
C280  
C281  
C282  
C283  
C284  
C285  
C286  
C287  
C288  
C289  
C290  
C291  
C292  
C293  
C294  
C295  
C296  
C297  
C298  
C299  
C300  
C301  
C302  
C303  
C304  
C305  
C306  
C307  
C308  
C309  
C310  
C311  
C312  
C313  
C314  
C315  
C316  
C317  
C318  
C319  
C320  
C321  
C322  
C323  
C324  
C325  
C326  
C327  
C328  
C329  
C330  
C331  
C332  
C333  
C334  
C335  
C336  
C337  
C338  
C339  
C340  
C341  
C342  
C343  
C344  
C345  
C346  
C347  
C348  
C349  
C350  
C351  
C352  
C353  
C354  
C355  
C356  
C357  
C358  
C359  
C360  
C361  
C362  
C363  
C364  
C365  
C366  
C367  
C368  
C369  
C370  
C371  
C372  
C373  
C374  
C375  
C376  
C377  
C378  
C379  
C380  
C381  
C382  
C383  
C384  
C385  
C386  
C387  
C388  
C389  
C390  
C391  
C392  
C393  
C394  
C395  
C396  
C397  
C398  
C399  
C400  
C401  
C402  
C403  
C404  
C405  
C406  
C407  
C408  
C409  
C410  
C411  
C412  
C413  
C414  
C415  
C416  
C417  
C418  
C419  
C420  
C421  
C422  
C423  
C424  
C425  
C426  
C427  
C428  
C429  
C430  
C431  
C432  
C433  
C434  
C435  
C436  
C437  
C438  
C439  
C440  
C441  
C442  
C443  
C444  
C445  
C446  
C447  
C448  
C449  
C450  
C451  
C452  
C453  
C454  
C455  
C456  
C457  
C458  
C459  
C460  
C461  
C462  
C463  
C464  
C465  
C466  
C467  
C468  
C469  
C470  
C471  
C472  
C473  
C474  
C475  
C476  
C477  
C478  
C479  
C480  
C481  
C482  
C483  
C484  
C485  
C486  
C487  
C488  
C489  
C490  
C491  
C492  
C493  
C494  
C495  
C496  
C497  
C498  
C499  
C500  
C501  
C502  
C503  
C504  
C505  
C506  
C507  
C508  
C509  
C510  
C511  
C512  
C513  
C514  
C515  
C516  
C517  
C518  
C519  
C520  
C521  
C522  
C523  
C524  
C525  
C526  
C527  
C528  
C529  
C530  
C531  
C532  
C533  
C534  
C535  
C536  
C537  
C538  
C539  
C540  
C541  
C542  
C543  
C544  
C545  
C546  
C547  
C548  
C549  
C550  
C551  
C552  
C553  
C554  
C555  
C556  
C557  
C558  
C559  
C560  
C561  
C562  
C563  
C564  
C565  
C566  
C567  
C568  
C569  
C570  
C571  
C572  
C573  
C574  
C575  
C576  
C577  
C578  
C579  
C580  
C581  
C582  
C583  
C584  
C585  
C586  
C587  
C588  
C589  
C590  
C591  
C592  
C593  
C594  
C595  
C596  
C597  
C598  
C599  
C600  
C601  
C602  
C603  
C604  
C605  
C606  
C607  
C608  
C609  
C610  
C611  
C612  
C613  
C614  
C615  
C616  
C617  
C618  
C619  
C620  
C621  
C622  
C623  
C624  
C625  
C626  
C627  
C628  
C629  
C630  
C631  
C632  
C633  
C634  
C635  
C636  
C637  
C638  
C639  
C640  
C641  
C642  
C643  
C644  
C645  
C646  
C647  
C648  
C649  
C650  
C651  
C652  
C653  
C654  
C655  
C656  
C657  
C658  
C659  
C660  
C661  
C662  
C663  
C664  
C665  
C666  
C667  
C668  
C669  
C670  
C671  
C672  
C673  
C674  
C675  
C676  
C677  
C678  
C679  
C680  
C681  
C682  
C683  
C684  
C685  
C686  
C687  
C688  
C689  
C690  
C691  
C692  
C693  
C694  
C695  
C696  
C697  
C698  
C699  
C700  
C701  
C702  
C703  
C704  
C705  
C706  
C707  
C708  
C709  
C710  
C711  
C712  
C713  
C714  
C715  
C716  
C717  
C718  
C719  
C720  
C721  
C722  
C723  
C724  
C725  
C726  
C727  
C728  
C729  
C730  
C731  
C732  
C733  
C734  
C735  
C736  
C737  
C738  
C739  
C740  
C741  
C742  
C743  
C744  
C745  
C746  
C747  
C748  
C749  
C750  
C751  
C752  
C753  
C754  
C755  
C756  
C757  
C758  
C759  
C760  
C761  
C762  
C763  
C764  
C765  
C766  
C767  
C768  
C769  
C770  
C771  
C772  
C773  
C774  
C775  
C776  
C777  
C778  
C779  
C780  
C781  
C782  
C783  
C784  
C785  
C786  
C787  
C788  
C789  
C790  
C791  
C792  
C793  
C794  
C795  
C796  
C797  
C798  
C799  
C800  
C801  
C802  
C803  
C804  
C805  
C806  
C807  
C808  
C809  
C810  
C811  
C812  
C813  
C814  
C815  
C816  
C817  
C818  
C819  
C820  
C821  
C822  
C823  
C824  
C825  
C826  
C827  
C828  
C829  
C830  
C831  
C832  
C833  
C834  
C835  
C836  
C837  
C838  
C839  
C840  
C841  
C842  
C843  
C844  
C845  
C846  
C847  
C848  
C849  
C850  
C851  
C852  
C853  
C854  
C855  
C856  
C857  
C858  
C859  
C860  
C861  
C862  
C863  
C864  
C865  
C866  
C867  
C868  
C869  
C870  
C871  
C872  
C873  
C874  
C875  
C876  
C877  
C878  
C879  
C880  
C881  
C882  
C883  
C884  
C885  
C886  
C887  
C888  
C889  
C890  
C891  
C892  
C893  
C894  
C895  
C896  
C897  
C898  
C899  
C900  
C901  
C902  
C903  
C904  
C905  
C906  
C907  
C908  
C909  
C910  
C911  
C912  
C913  
C914  
C915  
C916  
C917  
C918  
C919  
C920  
C921  
C922  
C923  
C924  
C925  
C926  
C927  
C928  
C929  
C930  
C931  
C932  
C933  
C934  
C935  
C936  
C937  
C938  
C939  
C940  
C941  
C942  
C943  
C944  
C945  
C946  
C947  
C948  
C949  
C950  
C951  
C952  
C953  
C954  
C955  
C956  
C957  
C958  
C959  
C960  
C961  
C962  
C963  
C964  
C965  
C966  
C967  
C968  
C969  
C970  
C971  
C972  
C973  
C974  
C975  
C976  
C977  
C978  
C979  
C980  
C981  
C982  
C983  
C984  
C985  
C986  
C987  
C988  
C989  
C990  
C991  
C992  
C993  
C994  
C995  
C996  
C997  
C998  
C999  
C1000

Db 8 PW--ETPAV--IDYKIMEQFGVKP--VDVLGDLKEZ-----HFFRRNIIIGHERFERI 57  
Qy 77 LDAYENKKEFYLYTGRGSPSEAMHVCHLIPFFFTKWLQ---DVFNVPLVLTQMTDDEKYLW 133  
Db 58 VDAIKNNKEFAVVGSMMPGSK-MHFCHKVVDLLKFKYQYKTDNINIP-----ADJEAYWA 112  
Qy 134 KDLTLDQAYGDV-ENAKDIIACGFDINKTFPSDLDYMGSMGSKYKVVVKI--OKHVTFN 191  
Db 113 RNSPFTTKELA-NEYITNYIALGLDPEKINYVLSQYQKV-----KDLALILSKTNWS 167  
Qy 192 QVKGIFGFTDSDCIGKISPPAIOAPSFNSPQIPDRDT--DIQCLIPCAIDQDDPYFRM 249  
Db 168 EMKAIYGFKEGTNIGHVFAPIVQVADIL---HPQLDENLSPEPKPVVVPVGVGDQPHIRJ 224  
Qy 250 TRDVAPR---IGYKPKALLHSTFFPALQCACTKMSASDPSSTFLDITAKQIKTKVKNKA 306  
Db 225 TRDIANRAKEKFIPIPSSTYHRPMTGLLG--KMSSSKPFETALFLDDEKTVKKKIFS-A 281  
Qy 307 FSGGRDTIEHRQFGG---NCDVDVSFMYLTTFLEDDDKLEQIRKDYTSGAMLTGELKXA 363  
Db 282 KTGGRETLLEHKKYGVPBECVVVELFLY--HLILDDKELARIYQKCRSGELTCGCKKM 339  
Qy 364 LIEVLQPLIAEHOARKEVTDEIVK 388  
Db 340 AYERVVEFLKDLKEKREQAKEIAVK 364

Search completed: August 24, 2004, 18:30:57  
Job time : 27 secs

Query Match 19.4%; Score 409.5; DB 1; Length 370;  
Best Local Similarity 30.9%; Pred. No. 7e-26;  
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
Qy 17 PWVVTSSAKGIDYDLIVRFGSKIKDELINIRATGQRPHPHLPFRGIFFSRDMNOV 76

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:42:41 ; Search time 117 Seconds  
(without alignments)  
1081.391 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SHEGPDATZABEDFVDPWTV.....VTDEIVKEFMTPKLSDFQ 401

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTEMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertibrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID     | Description         |
|------------|--------|-------------|--------|-----------|---------------------|
| 1          | 1957   | 92.5        | 475    | 11 Q9DC65 | Q9dc65 mus musculus |
| 2          | 1957   | 92.5        | 481    | 11 Q99J58 | Q99j58 mus musculus |
| 3          | 1953   | 92.3        | 475    | 11 Q80ZY4 | Q80zy4 mus musculus |
| 4          | 1820   | 86.0        | 475    | 13 Q7ZWT7 | Q7wt7 xenopus lae   |
| 5          | 1537   | 72.6        | 305    | 11 Q70184 | Q70184 cavia porce  |
| 6          | 1374.5 | 65.0        | 420    | 5 Q9U4Y0  | Q9u4y0 drosophila   |
| 7          | 1374.5 | 65.0        | 430    | 5 Q9U4Y1  | Q9u4y1 drosophila   |
| 8          | 1370.5 | 64.8        | 430    | 5 Q9VHG2  | Q9vhg2 drosophila   |
| 9          | 1323   | 62.5        | 417    | 5 Q9ULR2  | Q9ulr2 caenorhabdi  |
| 10         | 1301   | 61.5        | 402    | 10 Q9SKL5 | Q9sr15 arabidopsis  |
| 11         | 1209   | 57.1        | 491    | 3 Q8700C  | Q8700c neurospora   |
| 12         | 1026   | 48.5        | 632    | 5 Q8LDW3  | Q8ldw3 plasmodium   |
| 13         | 942    | 44.5        | 385    | 5 Q8SQY5  | Q8sqy5 encephalito  |
| 14         | 795.5  | 37.6        | 490    | 5 Q9ULF5  | Q9ulf5 leishmania   |
| 15         | 682    | 32.2        | 136    | 6 Q95295  | Q95295 sus scrofa   |
| 16         | 400.5  | 18.9        | 165    | 10 Q7XY51 | Q7xy51 griffithsia  |

|    |       |      |     |           |                    |
|----|-------|------|-----|-----------|--------------------|
| 17 | 310   | 14.7 | 111 | 5 Q95YL8  | Q95yl8 encephalito |
| 18 | 296.5 | 14.0 | 136 | 6 Q9TS88  | Q9ts88 bos taurus  |
| 19 | 294   | 13.9 | 157 | 5 Q9U533  | Q9u533 trypanosoma |
| 20 | 286   | 13.5 | 109 | 4 Q9UD15  | Q9ud15 homo sapien |
| 21 | 289.5 | 12.7 | 513 | 17 Q9HN83 | Q9hn83 halobacteri |
| 22 | 203.5 | 9.6  | 341 | 16 Q8E2J5 | Q8e2j5 streptococc |
| 23 | 203.5 | 9.6  | 341 | 16 Q8DWP7 | Q8dwp7 streptococc |
| 24 | 191   | 9.0  | 324 | 17 Q8TXZ2 | Q8txz2 methanopyru |
| 25 | 188   | 8.9  | 364 | 17 Q9YA64 | Q9ya64 aeropyrum p |
| 26 | 185.5 | 8.8  | 366 | 16 Q8ALC7 | Q8al07 bacteroides |
| 27 | 181.5 | 8.6  | 331 | 17 Q975Z1 | Q97921 thermoplasm |
| 28 | 181   | 8.6  | 340 | 16 Q8DRR1 | Q8drri streptococc |
| 29 | 178   | 8.4  | 340 | 16 Q87Q03 | Q87qg3 vibrio para |
| 30 | 174   | 8.2  | 334 | 16 Q832J8 | Q832j8 enterococcu |
| 31 | 171.5 | 8.1  | 351 | 10 P93018 | P93018 arabidopsis |
| 32 | 171.5 | 8.1  | 385 | 10 Q8S9T2 | Q8s9t2 arabidopsis |
| 33 | 170.5 | 8.1  | 376 | 5 Q7YYA0  | Q7yya0 cryptospori |
| 34 | 167   | 7.9  | 351 | 16 Q9RVD6 | Q9rvd6 deinococcus |
| 35 | 166.5 | 7.9  | 338 | 16 Q88ZB4 | Q88zb4 lactobacill |
| 36 | 164.5 | 7.8  | 377 | 5 Q86A90  | Q86a90 dictyosteli |
| 37 | 164.5 | 7.8  | 895 | 10 Q9SGN2 | Q9sgn2 arabidopsis |
| 38 | 158.5 | 7.5  | 356 | 16 Q83A61 | Q83a61 coxiella bu |
| 39 | 158.5 | 7.5  | 375 | 17 Q9V027 | Q9v027 pyrococcus  |
| 40 | 156   | 7.4  | 372 | 17 Q8Z1F7 | Q8zy07 pyrobaculum |
| 41 | 154.5 | 7.3  | 102 | 1 O07119  | O07119 halobacteri |
| 42 | 153.5 | 7.3  | 317 | 17 Q8TS11 | Q8ts11 methanosaer |
| 43 | 152.5 | 7.2  | 682 | 5 Q9N9B8  | Q9n9b8 leishmania  |
| 44 | 152   | 7.2  | 317 | 17 Q8PVK0 | Q8pvk0 methanosaer |
| 45 | 149.5 | 7.1  | 327 | 17 Q9HN62 | Q9hn62 halobacteri |

## ALIGNMENTS

RESULT 1

Q9DC65 Q9DC65 PRELIMINARY; PRT; 475 AA.

01-JUN-2001 (TRENBLrel. 17, Created)  
01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
Adult male lung cDNA, RIKEN full-length enriched library,  
clone:1200002C07, full insert sequence.  
WARS.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=C57BL/6J; TISSUE=Lung;  
MEDLINE=21085660; PubMed=11217851;  
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukumishi Y., Konno H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,  
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
Lyons P., Marchionni L., Mashima J., Mazzarelli C., Mommaerts P.,  
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
Hayashizaki Y.  
"Functional annotation of a full-length mouse cDNA collection."  
Nature 409:685-690(2001).  
EMBL; AK004541; BAB23357.1; -.

|    |  |
|----|--|
| DR | InterPro; IPR001412; tRNA-synt_I.                  |
| DR | InterPro; IPR002306; Ttp tRNA-synt_ib.             |
| DR | InterPro; IPR00738; WHEP-TRS.                      |
| DR | Pfam; PF00579; tRNA-synt_ib; 1.                    |
| DR | Pfam; PF00458; WHEP-TRS; 1.                        |
| DR | PRINTS; PR01039; TENASYNTHTRP.                     |
| DR | TIGRFAMs; TIGR00233; ttps; 1.                      |
| DR | PROSITE; PS00178; AA TRNA LIGASE_I; 1.             |
| DR | PROSITE; PS00762; WHEP TRS; 1.                     |
| KW | Aminoacyl-tRNA synthetase.                         |
| SQ | SEQUENCE 481 AA; 54325 MW; A754ELDDF58B2EF3 CRC64; |

  

|                           |  |                                       |
|---------------------------|--|---------------------------------------|
| Query Match               |  | 92.5%; Score 1957; DB 11; Length 481; |
| Best Local Similarity     |  | 91.5%; Pred. No. 6.8e-163;            |
| Matches 366; Conservative |  | 20; Mismatches 14; Indels 0; Gaps 0;  |

  

|    |     |  |     |
|----|-----|--|-----|
| QY | 2   | NHGPDATAEAEEDFVDPMVTQTSSAKGIDYDKLIVRGSSKIDKLINIERATGQRPHHF   | 61  |
| DB | 76  | NCDSDATKASEDFDPMVTVRTSSAKGIDYDKLIVQFGSKIDKLINIERATGQRPHRF    | 135 |
| QY | 62  | LRRGIFFSHRDMMQVLDAYENKKPFYLITGRGPSSEAMHVGHLIPFIETKWLQDVFNVEL | 121 |
| DB | 136 | LRRGIFFSHRDMMQLLDAYENKKPFYLITGRGPSSEAMELSHLVPFIETKWLQDVFNVEL | 195 |
| QY | 122 | VIOMTDEKYLKDILTDQAYGDVAENAKDIACGFINDTKTIFSDLDYMGSSGFYKNV     | 181 |
| DB | 196 | VIQMSDDEKYLKDLTLTEQAYSYYVENAKDIACGFINDTKTIFSDLEVMGSGPGFYRN   | 255 |
| QY | 182 | VKIQKHVFNQWGIFGTSDDCIGKISFPALQAAPSNFSFPQIFRDRDTIQCILPCA      | 241 |
| DB | 256 | VKIQKHVFNQWGIFGTSDDCIGKISFPAVQAAPSNFSFPKIFRDRDTIQCILPCA      | 315 |
| QY | 242 | DQDPYFRMTADVAPRICYPKPALLHSTFFFPALQAOTKVSASDPNSSIFLTDTAKOIKTK | 301 |
| DB | 316 | DQDPYFRMTADVAPRIHPKPALLHSTFFFPALQAOTKVSASDPNSSIFLTDTAKOIKTK  | 375 |
| QY | 302 | VNKHAFSGGRDTIEHRQFGNCVDVSYFWYLTFFLEDEEKLQIRKDYTSGLMTGLX      | 361 |
| DB | 376 | VNKHAFSGGRDTVEHRQFGNCVDVSYFWYLTFFLEDDELRLEQLRKDYTSGLMTGLX    | 435 |
| QY | 362 | KALIEVLQPLIAHQARKEVTDIVKEFMTPKRLSFDQ                         | 401 |
| DB | 436 | KTLLDIVLQPLIAHQARKAVTTETVKFEFTPRLSFHF                        | 475 |

  

|          |  |
|----------|--|
| RESULT 3 |  |
| ID       | Q80ZY4 PRELIMINARY; PRT; 475 AA.                                   |
| AC       | Q80ZY4;  |
| DT       | 01-JUN-2003 (TrEMBLrel. 24, Created)                               |
| DT       | 01-JUN-2003 (TrEMBLrel. 24, Last sequence update);                 |
| DE       | 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)                |
| DE       | Tryptophanyl-tRNA synthetase.                                      |
| OS       | Mus musculus (Mouse).  |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |
| OC       | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| RN       | NCBI_TaxID=10090;  |
| OX       | [1]  |
| RP       | SEQUENCE FROM N.A.   |
| RC       | STRAIN=C57BL/6; TISSUE=Brain;                                      |
| RA       | Straussberg R.;  |
| RL       | Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.            |
| RL       | EMBL; BC046232; AAH46232.1; -.                                     |
| DR       | GO; GO:0005524; F-ATP binding; IEA.                                |
| DR       | GO; GO:0004830; F-tryptophan-tRNA ligase activity; IEA.            |
| DR       | GO; GO:0006436; P-tryptophanyl-tRNA aminoacylation; IEA.           |
| DR       | InterPro; IPR002305; tRNA-synt_ib.                                 |
| DR       | InterPro; IPR001412; tRNA-synt_I.                                  |
| DR       | InterPro; IPR002305; Ttp tRNA-synt_ib.                             |
| DR       | InterPro; IPR000739; WHEP-TRS.                                     |
| DR       | Pfam; PF00579; tRNA-synt_ib; 1.                                    |
| DR       | Pfam; PF00458; WHEP-TRS; 1.  |
| DR       | PRINTS; PR01039; TENASYNTHTRP.                                     |



|  |  |   |     |
|--|--|---|-----|
| Q1   | 157  | FDINKTTFISDLIDYMGMSGFFKYNVVKIKQHVTEHQVKIGFQGTSDCTGKISFPAIQAA  | 218 |
| D1   | 61   | FDINKTTFISDLIDYMGMSGFFKYNVVKIKQHVTEHQVKIGFQGTSDCTGKISFPAIQAA  | 120 |
| Q1   | 217  | PSFSNSFPQIFRDRTDIOCLIPCAIDQDPYFRMTDVAIPRIGYKPKALHSTFFFPALOGA  | 276 |
| D1   | 121  | PSFSNSFPQIFRDRTDIOCLIPCAIDQDPYFRMTDVAIPRIGYKPKALHSTFFFPALOGA  | 180 |
| Q1   | 277  | QTKMSASDPNSISIFLDTAKQIKTKVKNHAFSGGSDTIEHHRQFGNGCDVDVSFMYLTFF  | 336 |
| D1   | 181  | QTKMSASDPNSISIFLDTAKQIKTKVKNHAFSGGSDTIEHHRQFGNGCDVDVSFMYLTFF  | 240 |
| Q1   | 337  | LEDDKLEQIRKDYTSGLMTGELKALIEVLQPLIAEHQARKEVTDIEIVKEFMTPRKL     | 396 |
| D1   | 241  | LEDDKLEQIRKDYTSGLMTGELKALIEVLQPLIAEHQARKEVTDIEIVKEFMTPRKL     | 300 |
| Q1   | 397  | SFDFQ 401   |     |
| D1   | 301  | SFDFQ 305   |     |
| RESULT 6   |  |   |     |
| Q1   | Q9U4Y0   | PRELIMINARY; PRT; 420 AA.                                     |     |
| AC   | Q9U4Y0;  |   |     |
| DT   | 01-MAY-2000 (TREMBlurel. 13, Created)                                |   |     |
| DT   | 01-MAY-2000 (TREMBlurel. 13, Last sequence update)                   |   |     |
| DT   | 01-OCT-2003 (TREMBlurel. 25, Last annotation update)                 |   |     |
| DE   | Tryptophanyl-tRNA synthetase (Fragment).                             |   |     |
| GN   | AATS-TRP OR CG9735.  |   |     |
| OS   | Drosophila melanogaster (Fruit fly).                                 |   |     |
| OC   | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prterygota;       |   |     |
| OC   | Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;           |   |     |
| OC   | Eohydridea; Drosophilidae; Drosophila.                               |   |     |
| OX   | NCBI_TaxID=7227;   |   |     |
| RN   | [1]  |   |     |
| RP   | SEQUENCE FROM N.A.   |   |     |
| RA   | MEDLINE=99250164; PubMed=10233165;                                   |   |     |
| RX   | Seshaiah P., Andrew D.J.;  |   |     |
| RT   | "WRS-85D; A tryptophanyl-tRNA synthetase expressed to high levels in |   |     |
| RT   | the developing Drosophila salivary gland."                           |   |     |
| RT   | Mol. Biol. Cell 10:1595-1608(1999).                                  |   |     |
| DR   | EMBL; AF125157; AAF20167.1; -;                                       |   |     |
| DR   | FlyBase; FBgn0010803; Aats-tp.                                       |   |     |
| DR   | GO; GO:0005524; P-ATP binding; IEA.                                  |   |     |
| DR   | GO; GO:0006436; P-tryptophanyl-tRNA ligase activity; IEA.            |   |     |
| DR   | InterPro; IPR002305; tRNA-synt 1b.                                   |   |     |
| DF   | InterPro; IPR014112; tRNA-synt 1.                                    |   |     |
| DF   | Pfam; PF00579; tRNA-synt_1b.   |   |     |
| DF   | PRINTS; PR01039; TRNASYNTHTRP.                                       |   |     |
| DF   | TIGRfam; TIGR00233; trps; 1.   |   |     |
| DF   | PROSITE; PS00178; AA TRNA LIGASE_I; 1.                               |   |     |
| DF   | Aminoacyl-tRNA synthetase.   |   |     |
| FT   | NON TER  |   |     |
| K1   | SEQUENCE 420 AA; 46975 MW; 4AF7086AF426AEB6 CRC64;                   |   |     |
| SC   |  |   |     |
| Query Match  |  |   |     |
| Best Local Similarity 65.0%; Score 1374.5; DB 5; Length 420;   |  |   |     |
| Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1; |  |   |     |
| Q1   | 7  | ATEAEDFDVDPWTVQTSSAKGIDYDKLIVRPGSSKIDKELINRIERATGCRPHFLRGI    | 66  |
| D1   | 28   | ATAETEDVDPWTVQTNVASSNDAGVDYDKLIVRPGSSKIDKELINRIERATGCRPHFLRGI | 87  |
| Q1   | 67   | FFSHRDMNVLDAYENKPFYLYTCRGPSSSEAWYCHLTPFTKWLQDVFNVLIVMT        | 126 |
| D1   | 88   | FFSHRDLHLTLTRQCKPFYLYTCRGPSSGSLRHVHLVPEFTMTKWLQCFDVLIVLT      | 147 |
| Q1   | 127  | DDEKYLKDLTLDAQYDAVENAKDIIACGPDINKTFFPSDLIDYMGMSGFFKYNVVKIKQ   | 186 |



QY 67 FFSHRDMMQVLDAYENKPPFLYLTGRGSSSEAMVGHLPFIPTFKWLOQVFNVLVIOMT 126  
 DB 98 FFSHRDLHTLTREQGGKPPFLYLTGRGSSSLVGHLPFIMTKWLOQVFNVLVIOMT 157  
 QY 127 DDEKYLWKOLTLDOAYGDVAENAKDIACGDFINKTIFPSDLDMGSSGGYKKNVVKIQ 186  
 DB 158 DDEKTLWKOLKVEDAIKLGRENKDIIVAIGDFVNTKIFNNLEFVKGKPMYQMIIRIQ 217  
 QY 187 HVTFNQVKGIFGTDSDCIKISPAOAPSFNSPFIQFRDRTDIQCLIPCAIDQDPY 246  
 DB 218 CVTFNQVKGIFGTDSDCIKISPAOAPSFNSPFIQFRDRTDIQCLIPCAIDQDPY 276  
 QY 247 FMTRDVAIPRIGVKKPALLHSTFPPALOGATKYSASDPNSIFLTTAKQIKTKVAKHA 306  
 DB 277 FMTRDVAIPRIGVKKPALLHSTFPPALOGATKYSASDPNSIFLTTAKQIKTKVAKHA 336  
 QY 307 FSGGRDTIEHROFGGNCDDVDVFMVLTFFLEDECKLEQIRKDYTSAMLTGELKXALIE 366  
 DB 337 FSGGRDTIEHROFGGNCDDVDVFMVLTFFLEDECKLEQIRKDYTSAMLTGELKXALIE 396  
 QY 367 VLOPLIAHQARKEVTDDEIVKEMTPRKUSP 398  
 DB 397 TLTPIVEQHQARKLTIDVLDKYFELRPLKF 428

RESULT 8  
 OSVHG2  
 Q9VHG2 PRELIMINARY; PRT; 430 AA.  
 Q9VHG2  
 01-MAY-2000 (TREMBlrel. 13, Created)  
 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 CG9735-PA.  
 GN  
 AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 GN Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 GN Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 GN Ephydroidea; Drosophilidae; Drosophila.  
 GN NCBI\_TaxID=7227;  
 RN [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=20196006; PubMed=10731132;  
 ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.H., Blazey R.G., Champagne M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 Surtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cterry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 Foster C., Gabrielian A.E., Garg N.J., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.H., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,  
 Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattel J., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson S.D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Poilard J., Puri V., Reese V.G.,  
 Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*."  
 Science 287:2185-2195(2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 Banzon J.J., An H., Baldwin D., Beeson K.V., Busan D.A.,  
 Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 Dodson K., Dorsett V., Douc L.E., Doyle C., Dresnek D., Farfan D.,  
 Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 Ibegwan C., Jallali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
 McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 "Sequencing of *Drosophila melanogaster* genome."  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 "Annotation of *Drosophila melanogaster* genome."  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RN SEQUENCE FROM N.A.  
 RA FlyBase;  
 RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN EMBL; AE003682; AAF54352.1; --  
 DR FlyBase; FBgn0010803; Aacs-trp.  
 DR GO; GO:0005524; F-ATP binding; IEA.  
 DR GO; GO:0004630; F-tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P-tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1.  
 DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAM; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 DR SEQUENCE 430 AA; 47971 MW; 2938ESEC69E979F CRC64;

Query Match 64.8%; Score 1370.5; DB 5; Length 430;  
 Best Local Similarity 64.5%; Pred. No. 1.6e-111;  
 Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;  
 QY 7 ATEAEEDFVDPVTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQPHHFLRGI 66  
 DB 38 ATAPTEDVDVFNWASSNDAGVDYDKLIVRFGSSKIDKELINRIERATGQPHHFLRGI 97  
 QY 67 FFSHRDMMQVLDAYENKPPFLYLTGRGSSSEAMVGHLPFIPTFKWLOQVFNVLVIOMT 126  
 DB 98 FFSHRDLHTLTREQGGKPPFLYLTGRGSSSLVGHLPFIMTKWLOQVFNVLVIOMT 157  
 QY 127 DDEKYLWKOLTLDOAYGDVAENAKDIACGDFINKTIFPSDLDMGSSGGYKKNVVKIQ 186  
 DB 158 DDEKTLWKOLKVEDAIKLGRENKDIIVAIGDFVNTKIFNNLEFVKGKPMYQMIIRIQ 217

```

2Y 187 HVTFNQVKGIFGFTSDSDICIGKISFPALQAAAPSFNSPQIFRDRDTIOCLIPCAIDQDPY 246
2b 218 CVTFNQVKGIFGFTSDSDIIGKIGFPAAQAAAPASSTPFIFGNR-KVHCLIPCAIDQDPY 276
2Y 247 FMTTRDVAAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKXHA 306
2b 277 FMTTRDVAAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKXHA 336
2Y 307 FSGGRDTIEHRQFGGNCVDVNSFMVLTFFLEDDDKLEQIRKDYTSAMLTGELKKALIE 366
2b 337 FSGGRVSVVEHRKLGVPDVSVYQLLKFFLEDDAKLEEVVAVYSGEMLTGHIKKLAVE 396
2Y 367 VLQPLTAEHQARKEVTDEIVKPFMTPRKLSF 398
2b 397 TLTPIVEHQAAARKLITDEVLKDFELRPLKF 428

RESULT 9
19UR2 PRELIMINARY; PRT; 417 AA.
Q9UR2;
Q9UR15 PRELIMINARY; PRT; 402 AA.
AC Q9SR15;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase.
GN F7018.7 OR AT3G04600.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito X., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White C., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence.";
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Full Length cDNA Clones.";
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale C.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Open Reading Frame (ORF) Clones.";
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA AC011437; AAF04890.1; -
RA EMBL; AY080709; AAL85027.1; -
RA EMBL; AY117275; AAM51350.1; -
RA GO; GO:0005224; F:ATP binding; IEA.
RA GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
RA GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.

```

```

QY 124 QMTDEKYLKOLTLDOAYGDAVENAKDIACGFDINKTIFISDLDMGSSGYKNVVK 183
Db 143 QMTDEKFLKMKVDEAKKVAENKMDIISVGFDPTKTFIFANNFY--XCPPEYENIVK 200
QY 184 IQKHVTFNQVKGIFGFTSDSDICIGKISFPALQAAAPSFNSPQIFRDRDTIOCLIPCAIDQ 243
Db 201 IKWVNTQNAALFGTTPEDCLGAAPFAVEAAPCFASFPQIFGRNDIPCLIPCAIDQ 260
QY 244 DPFMTTRDVAAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKX 303
Db 261 DPFMTTRDVAAPRIGYKPKALHSTFFPALQGAQTKMSADPNSSIFLTDTAQIKTKKX 320
QY 304 KHAFFSGGRDTIEHRQFGGNCVDVNSFMVLTFFLEDDDKLEQIRKDYTSAMLTGELKK 363
Db 321 KYAFSGGQOTVQSHRKGKGCNDVDSYQFLRFFLDDEKLAIE-RENYTKGEMLSGELKAL 380
QY 364 LIEVLQPLTAEHQARKEVTDEIVKPFMTPRKLSF 400
Db 381 ATQKVCQEVLEMQERKLVDTETVEEFVKRPLAYKY 417

```

## RESULT 10

```

Q9SR15
AC Q9SR15;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase.
GN F7018.7 OR AT3G04600.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito X., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White C., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence.";
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Full Length cDNA Clones.";
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale C.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Open Reading Frame (ORF) Clones.";
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA AC011437; AAF04890.1; -
RA EMBL; AY080709; AAL85027.1; -
RA EMBL; AY117275; AAM51350.1; -
RA GO; GO:0005224; F:ATP binding; IEA.
RA GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.
RA GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.

```

```

Query Match 62.5%; Score 1323; DB 5; Length 417;
Best Local Similarity 62.0%; Pred. No. 2.2e-107;
Matches 246; Conservative 63; Mismatches 86; Indels 2; Gaps 1;

4 GPDATAREADVPWTVQTSAGKIDYDKLIVRPGSSKIDKELINLRIERATGQPHFIR 63
23 GGGVQDEEDRVTFWETTTKATGIDKDKLVKPGCKLDEIIRFERYTGHKASPLR 82
64 RGIFFSHRDMQVLDAYENKKPFYLYTGRGSSAMHWGHLIPFIPTKWLQDVNFVLVI 123
83 RGMFFAHRDLTALDRKQKGPFLYTGRCASSGLHLGVLPFIPTKWLQDVNFVLVI 142

```

DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1b.  
 DR InterPro; IPR002306; Trp-tRNA-synt 1b.  
 DR Pfam; PF00579; tRNA-synt 1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMS; TIGR00233; tips; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75E5D6CD15 CRC64;  
 Query Match 61.5%; Score 1301; DB 10; Length 402;  
 Best Local Similarity 60.3%; Pred. No. 1.8e-105;  
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;  
 6 DATEAE--EEFVDPWTQTSSAKGIDYDKLVIRFGSSKIDKELINRIERATGQPHFLR 63  
 7 DEREAESEQVNPWEYSADGGGIDYDKLVIRFGSSKIDKELINRIERATGQPHFLR 66  
 64 RGIFFSHRDNMQVLDAYENKKPFYLTGRGSSSEAMVGHILIPFTFKWLQDVNVLVI 123  
 67 RSVFFAHRDNEILDATERGDKFYLTGRGSSSEALHGLHILIPFTFKYLQEAFFVLVI 126  
 124 QMTDDEKYLWKDLTQDAYDAVENAKDIIACGFDINKTIFISDLVGMSSGFGYKXWK 183  
 127 QLTDDEKSIWNLSEVRSLARENKDIILACGFDVTKTIFISDFYVG--GAFYKXWK 184  
 184 IQKHVTFNQVKGIEGFTDSCIGKISPAIQAPSPNSPQIPFRDRTDIQCLIPCAIDQ 243  
 185 VGCVTNLKAWG:FGSGEDPIAKLSPPVQAVSPSPSPHLPFGKDNRLCLIPCAIDQ 244  
 244 DYFERTRDVAIRIGYKPKALLHSTFPALQGAQTKMSASDPNSSIFLDTAQIKTKVN 303  
 245 DYFERTRDVAIRIGYKPKALLHSTFPALQGAQTKMSASDPNSSIFLDTAQIKTKVN 304  
 304 KNAFGSGRTIEHROFGGNCVDVSMYLTFFLEDKLEQIRKDYTSGLMTGELKKA 363  
 305 RYAFSGGQDSIEKHRELGNLEVDIPVKYLSFFLEDDSELEHKKEGGRMTGVEKKR 364  
 364 LLEVLQPLAEHQARKEVTDIVKEFMPTRKLSDFQ 401  
 365 LFEVLTEIVERHRRRAAATVDMVDAMFAVRPLFSMFE 402

RESULT 11  
 Q870UO PRELIMINARY; PRT; 491 AA.  
 ID Q870UO  
 AC Q870UO  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Probable tryptophan--tRNA ligase.  
 GN BLH7.080.  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 NCBI\_TaxID=5141;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 [2]  
 SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL EMBL; BX294092; CAD71247.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR001412; tRNA-synt 1b.  
 DR InterPro; IPR002306; Trp-tRNA-synt 1b.

DR Pfam; PF00579; tRNA-synt 1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMS; TIGR00233; tips; 1.  
 DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Ligase.  
 SQ SEQUENCE 491 AA; 55393 MW; C2833C847A66C1D9 CRC64;  
 Query Match 57.1%; Score 1209; DB 3; Length 491;  
 Best Local Similarity 58.6%; Pred. No. 2.7e-97;  
 Matches 243; Conservative 53; Mismatches 97; Indels 22; Gaps 6;  
 5 PDAT--AEEDFDPWTQV-----TSSAKGIDYDKLVIRFGSSKIDKELINRIERATGQ 56  
 16 ELAIEKHAASQTVDPYNNQGVGSDGVVKAIDYNNKLEIEFGTKKIDQALLDLRLERVTK 75  
 57 RPHEFLRGIEFHSRDNMQVLDAYENKKPFYLTGRGSSSEAMVGHILIPFTFKWLQDV 116  
 76 KPHFLRGIVFHSRHELELDIERYEKGEFFYLTGRGSSSDSVHVGHTIPEFTKWLQDT 135  
 117 FNVLVIQMTDDEKYLWKDLTQDAYDAVENAKDIIACGFDINKTIFISDLVGMSS 175  
 136 FDVLVIQMTDDEKYLWKDLTQDAYDAVENAKDIIACGFDINKTIFISDLVGMSS 193  
 176 GFYKXVVKIQKHVTFNQVKGIEGFTDSCIGKISPAIQAPSPNSPQIPFRDRTDI 231  
 194 AFYKXVVKLSKHTLNOARALFGFNDSTN--GRIHFGSLQGASSWASSFPHFGEDESKTV 253  
 232 DIQCLIPCAIDQDQPFYFMTDRVAPR-----ISYKPKALLHSTFPALQGAQTKMSASD 284  
 254 AIPCLIPCAIDQDQPFYFMTDRVAPR-----ISYKPKALLHSTFPALQGAQTKMSASD 313  
 285 PNSSIFLDTAQIKTKVNKHAPSGGRTTIEHROFGGNCVDVSMYLTFFLEDKLE 344  
 314 DESAIFMDETPNQIKNKINKTAFSGGKVTVEEHEKGGDTNVDVAYQTLRFFLEDDEE 373  
 345 OIRKDYTSGLMTGELKKAIEVLQPLAEHQARKEVTDIVKEFMPTRKLSDFQ 399  
 374 RIRVAVESGDMLTGDLKAICIKELQTVVAAQFERRAKVDDEAVKLFMTTRPLKWN 428  
 RESULT 12  
 Q81DW3 PRELIMINARY; PRT; 632 AA.  
 ID Q81DW3  
 AC Q81DW3  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tryptophan--tRNA ligase, putative (EC 6.1.1.2).  
 GN PF13\_0205.  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 NCBI\_TaxID=36329;  
 [1]  
 SEQUENCE FROM N.A.  
 RA Harris B., Lemard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriaman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrall B.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844509; CAD52506.1;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004830; F:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt 1b.  
 DR InterPro; IPR002306; Trp-tRNA-synt 1b.  
 DR Pfam; PF00579; tRNA-synt 1b; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR TIGRFAMS; TIGR00233; tips; 1.  
 KW Ligase.  
 SQ SEQUENCE 632 AA; 73099 MW; B3E9DE08A4BF0C51 CRC64;  
 Query Match 48.5%; Score 1026; DB 5; Length 632;  
 Best Local Similarity 50.1%; Pred. No. 4.4e-81;

Matches 203; Conservative 66; Mismatches 114; Indels 22; Gaps 5;

QY 9 BAEDFDPWTVQTSAGIDYDKLIVRFGSKIDKELINRIERATGQPHHFLRRGIPF 68  
 DI 227 ERSKD-VTFWDVNNINERGINYNKLIKFGGSKITENHIKRIEKLNSKAHFIIRRGIPF 285  
 QY 69 SHRDNOVLDAVENKPKLYLTGRGSPSEAMHVGLHIFPIFKWLODVNPLVLTQMTDD 128  
 DI 286 SHRDDELNLYEHOHKFYIYTGGRGSSLSMHLGLHIFPFCKYLQEAFAFVPLVLTQSD 345  
 QY 129 EKLW-KDLTLDQAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKXNVKIOKH 187  
 DI 346 EKLWQNVNLYEINTLNTNENKDIISVGLPELTFIKNEYAGY---LPTVLISHK 402  
 QY 188 VTFNQVGIIFTSDCIKISFPAIOAAPSFSNFPQIFRDRTDIQLIPCAIDQDPYF 247  
 DI 403 TTLNQSMNVFNGFNSDNIGKISYFSFQIAPCFSCFPNPL--GKNIPCLVPGIDQDPYF 460  
 QY 248 RMTDVAAPRIGYPKPKALLHSTFFPALOGAQTMSASDP-----NSSIFLT 292  
 DI 461 RLSRDIAVKALHKPVVHSVFMCLQGVNSKMSSTKKKDDNGKSNSTFDHNSVIFLT 520  
 QY 293 DTAQIKTKVNGHAFSGGRDTIEHRQFGNCDDVDVSMYLTFFLEDDKLEQIRKDYTS 352  
 DI 521 DTPEQIKNKINKYAFSGGGTTIQEHREKGNLKDYSQYLYRLLEDGKLEIGEKYKK 580  
 QY 353 GAMLTELKALIEVLOPLTAHQARKEVTDEIVKSFMTPRKLS 397  
 DI 581 GEMLSGEIKLILDLVLTDLVHLQEKKSLTDEISYFFDPNPKS 625

RESULT 13  
 ID Q8SQY5 PRELIMINARY; PRT; 385 AA.  
 AC Q8SQY5;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tryptophanyl tRNA synthetase.  
 GN ECUL1.0530.  
 OS Eucaryote.  
 OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.  
 OX NCBI\_TaxID=6035;  
 RN [1]  
 RC STRAIN=GB-M1;  
 RA Genoscope;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RS STRAIN=GB-M1;  
 RX MEDLINE=21576510; PubMed=11719806;  
 RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat P.,  
 RA Frensier G., Barbe V., Peyretailade B., Brottier P., Wincker P.,  
 RA Delbac P., El Alaoui H., Peyret P., Saurin W., Gouy M.,  
 RA Weissenbach J., Vivares C.P.;  
 RT "Genome sequence and gene compaction of the eukaryote parasite  
 RT Encephalitozoon cuniculi";  
 RL Nature 414:450-453(2001).  
 RL EMBL; AL530450; CAD25963.1; --  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR TIGRFAMs; TIGR00233; trps; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 SQ SEQUENCE 385 AA; 4418 MW; 28653AC95FCC859 CRC64;

Query Match 44.5%; Score 942; DB 5; Length 385;

Best Local Similarity 46.1%; Pred. No. 5e-74;  
 Matches 179; Conservative 76; Mismatches 123; Indels 10; Gaps 3;

QY 12 EDFVDPWTVQTSAGIDYDKLIVRFGSKIDKELINRIERATGQPHHFLRRGIPF 67  
 DI 3 EQRITFDVWVSTDEVPAIDYDKIINQFGCEKFNQALADLEKLSGKPAHVFRRGIV 62  
 QY 68 FSHRDNOVLDAVENKPKLYLTGRGSPSEAMHVGLHIFPIFKWLODVNPLVLTQMTD 127  
 DI 63 FAHRDENLLEIANNRPFYITGRGSSKTMHGHITPFLLCKYMQDAFKRLVLTQTD 122  
 QY 128 DEKYLMDLTDQAYGDAVENAKDIIACGFDINKTFISDLDYMGSSGFYKXNVKIOKH 187  
 DI 123 DEKFLMSMRLEADAMAYGRENKOIIVALGDFPKLTYIESNVE--ASHFENIKIKST 179  
 QY 188 VTFNQVGIIFTSDCIKISFPAIOAAPSFSNFPQIFRDRTDIQLIPCAIDQDPYF 247  
 DI 180 ININEATKVFDFSSNIGQVGFPAKEIAPCFSSSFRPIGK---GAMCLVPAAVDQDPFF 236  
 QY 248 RMTDVAAPRIGYPKPKALLHSTFFPALOGAQTMSASDPNSIFLTDATAKQIKTKVNGHAF 307  
 DI 237 RLARDKAKALGEKPKSPSIYVLSLLPDLKGVNRMKMSASDPNSIYLLDCAQPTIRKKTIAAY 296  
 QY 308 SGRDRTIEHRQFGNCDDVDVSMYLTFFLEDDKLEQIRKDYTS GAMLTELKALIEV 367  
 DI 297 SGRKRTLEHREKGGIDVDVPEVLYXVFLDDQDELEKYRSGYIKGEITSKEMKEKCVVV 356  
 QY 368 LOPLIAEHQARKEVTCEIVKEFMTPRK 395  
 DI 357 IQEFVSRYQESRKRVTDLLEAFIDINK 384

RESULT 14  
 ID Q9UIF5 PRELIMINARY; PRT; 430 AA.  
 AC Q9UIF5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Tryptophanyl tRNA synthetase.  
 GN LI063.04.  
 OS Leishmania major.  
 OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RC STRAIN=FROM N.A.  
 RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,  
 RA Barrell B.G.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RS STRAIN=Friedlin;  
 RX MEDLINE=98146435; PubMed=9477341;  
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 RA Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome";  
 RL Genome Res. 8:135-145(1998).  
 DR EMBL; AL121862; CAB58393.1; --  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004830; P:tryptophan-tRNA ligase activity; IEA.  
 DR GO; GO:0006436; P:tryptophanyl-tRNA aminoacylation; IEA.  
 DR InterPro; IPR002305; tRNA-synt\_1b.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR InterPro; IPR002306; Trp tRNA-synt\_1b.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS; PR01039; TRNASYNTHTRP.  
 DR PROSITE; PS00178; AA TRNA LIGASE I; 1.  
 DR Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 430 AA; 54904 MW; 172C68622C51D3D3D CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;  
 Best Local Similarity 38.1%; Pred. No. 5e-61;



GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 Compugen Ltd.  
 OM protein - protein search, using sw model  
 Run on: August 24, 2004, 17:47:41 : Search time 32 Seconds  
 (without alignments)  
 646,938 Million cell updates/sec  
 Title: US-09-813-718-10\_COPY\_71\_471  
 Perfect score: 2116  
 Sequence: 1 SNHGPDATCAEEDFVDPTV.....VTDEIVKEFMPRLKLSDFDQ 401  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | % Match | Length | DB ID | Description          |
|------------|--------|---------|--------|-------|----------------------|
| 1          | 1218.5 | 57.6    | 424    | 3     | US-08-876-885-26     |
| 2          | 195    | 9.2     | 338    | 4     | US-09-328-352-4201   |
| 3          | 186    | 8.8     | 348    | 4     | US-09-107-532A-5765  |
| 4          | 185.5  | 8.8     | 341    | 3     | US-08-928-100-2      |
| 5          | 185.5  | 8.8     | 341    | 4     | US-09-492-581-2      |
| 6          | 185.5  | 8.8     | 341    | 4     | US-09-425-666-2      |
| 7          | 174    | 8.2     | 335    | 4     | US-09-134-000C-5784  |
| 8          | 164.5  | 7.8     | 409    | 2     | US-08-743-130A-39    |
| 9          | 162.5  | 7.7     | 409    | 2     | US-08-743-130A-2     |
| 10         | 157.5  | 7.4     | 385    | 4     | US-09-489-039A-8660  |
| 11         | 152    | 7.2     | 344    | 4     | US-09-198-452A-857   |
| 12         | 132.5  | 6.3     | 388    | 2     | US-08-705-868-4      |
| 13         | 132.5  | 6.3     | 388    | 3     | US-09-123-615-4      |
| 14         | 123    | 5.8     | 356    | 4     | US-09-543-681A-4257  |
| 15         | 118    | 5.6     | 349    | 4     | US-09-543-681A-6769  |
| 16         | 118    | 5.6     | 426    | 4     | US-09-107-532A-4046  |
| 17         | 113.5  | 5.4     | 418    | 3     | US-08-855-910-11     |
| 18         | 113.5  | 5.4     | 433    | 4     | US-09-134-000C-3686  |
| 19         | 109    | 5.2     | 377    | 3     | US-09-352-990-28     |
| 20         | 108    | 5.1     | 197    | 2     | US-08-923-867-2      |
| 21         | 108    | 5.1     | 197    | 3     | US-08-928-100-4      |
| 22         | 108    | 5.1     | 197    | 3     | US-09-183-134-2      |
| 23         | 108    | 5.1     | 197    | 4     | US-09-492-581-4      |
| 24         | 108    | 5.1     | 197    | 4     | US-09-425-666-4      |
| 25         | 104.5  | 4.9     | 370    | 2     | US-08-415-593-45     |
| 26         | 100.5  | 4.7     | 427    | 4     | US-09-134-001C-5141  |
| 27         | 97.5   | 4.6     | 335    | 4     | US-09-489-039A-10919 |

|    |      |     |      |   |                     |                   |
|----|------|-----|------|---|---------------------|-------------------|
| 28 | 95   | 4.5 | 418  | 3 | US-08-844-054-2     | Sequence 2, Appli |
| 29 | 95   | 4.5 | 418  | 3 | US-09-347-333-2     | Sequence 2, Appli |
| 30 | 95   | 4.5 | 464  | 4 | US-09-134-001C-4701 | Sequence 4701, Ap |
| 31 | 94.5 | 4.5 | 877  | 2 | US-08-907-166-8     | Sequence 8, Appli |
| 32 | 94.5 | 4.5 | 877  | 4 | US-09-391-340-8     | Sequence 8, Appli |
| 33 | 92.5 | 4.4 | 434  | 4 | US-09-543-681A-4577 | Sequence 4577, Ap |
| 34 | 91.5 | 4.3 | 454  | 4 | US-09-459-039A-9702 | Sequence 9702, Ap |
| 35 | 91   | 4.3 | 892  | 4 | US-09-585-858-2     | Sequence 2, Appli |
| 36 | 91   | 4.3 | 892  | 4 | US-09-585-858-2     | Sequence 2, Appli |
| 37 | 89.5 | 4.2 | 344  | 3 | US-09-393-554-2     | Sequence 3, Appli |
| 38 | 89   | 4.2 | 1467 | 4 | US-09-134-000C-6740 | Sequence 6740, Ap |
| 39 | 88.5 | 4.2 | 501  | 4 | US-09-157-257-8     | Sequence 8, Appli |
| 40 | 88.5 | 4.2 | 898  | 1 | US-08-465-995A-4    | Sequence 4, Appli |
| 41 | 88.5 | 4.2 | 898  | 2 | US-08-465-995A-4    | Sequence 4, Appli |
| 42 | 88.5 | 4.2 | 898  | 2 | US-08-966-145-4     | Sequence 4, Appli |
| 43 | 88.5 | 4.2 | 920  | 1 | US-08-101-593-4     | Sequence 3, Appli |
| 44 | 88   | 4.2 | 606  | 2 | US-08-883-534-3     | Sequence 3, Appli |
| 45 | 88   | 4.2 | 606  | 3 | US-09-204-764-3     | Sequence 3, Appli |

## ALIGNMENTS

## RESULT 1

US-08-876-885-26  
 ; Sequence 26, Application US/08876885  
 ; Patent No. 6174713

; GENERAL INFORMATION:  
 ; APPLICANT: Shen, Xiaoyu  
 ; APPLICANT: Homan, Fariba

; TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPHANYL-TRNA  
 ; TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING

; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive  
 ; CITY: Lexington  
 ; STATE: MA  
 ; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/876,885

; FILING DATE: 16-JUN-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brook, David E.

; REGISTRATION NUMBER: 22,592  
 ; REFERENCE/DOCKET NUMBER: CPI97-02  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781) 861-6240

; TELEFAX: (781) 861-9540  
 ; INFORMATION FOR SEQ ID NO: 26:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 424 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear

US-08-876-885-26

Query Match 57.6%; Score 1218.5; DB 3; Length 424;  
 Best Local Similarity 57.1%; Pred. No. 4, 1e-128;  
 Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 8 TEAEEDFVDPTVQ-----TSSAKGIDYDKLIVFGSKDKKELINIERATGQRPHFLR 63

Db 23 TEESEQKITPWEVGAVDVKMSGIDYDKLISQGTGKHITETLRFKQVGTGEPHFLK 72

```

QY 64 RGIFFSRDMQVLDAYENKKPFVLYTGRGSPSEAMVGHLPFIPTKWLQDVFNWPLVI 123
DB 73 RGVPFSQDRLDILDYEHGEPFLYTGSPSSMELGHMWPFIFTKLQVDFWPLVI 132
QY 124 QMTDDKYLWK-DLTLDQAVGDAVENAKDIIACGFCJINKTFIFSDLDYMGSSGFYKVV 182
DB 133 ELTDDKFLFKHJLTDVKGFAENAKDIIAVGNFNTFIFSDLYQYMG--GAFYENVV 190
QY 183 KIQKHVFNQVKGIFGFTSDCIGKISFPFAIQAPSPNSFPQIFRDRTDIQCILPCAID 242
DB 191 RTSRQITTTAKAVGFTSDCIGKIHFAISQIATAPPSSFDVLGLPPTKPLPCAID 250
QY 243 QDPYFRTVDVAPRIGYPKAPLHSTFFPALQAGATMSASDPSIFLTDIAKQIKTV 302
DB 251 QDPYFVCRVDVADKRFKFAIHAFFPALQAGATMSASDPSIFLTDIAKQIKTV 310
QY 303 NKHAFSGGRDITIEHRQFGNCDVDSVYMLTFFLEDDDKLEQIRKDYTSGLMTGELKK 362
DB 311 NKYAFSGGRATAEHRHGLGNPEVDVAFQLSFFSYDDKLAQLEQYRKGELTSEMCK 370
QY 363 ALIEVLQPLIAEHQARKKEVTDBIVKFEFTPRKLSF 398
DB 371 ECIITVLEFVSAYQERRSKVDQVVEKPKHLPV 406

RESULT 2
US-09-328-352-4201
; Sequence 4201, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4201
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US 09-328-352-4201

Query Match      9.2%; Score 195; DB 4; Length 338;
Best Local Similarity 25.9%; Pred. No. 4e-13;
Matches 84; Conservative 56; Mismatches 138; Indels 46; Gaps 13;
QY 82 NKKPFVLYTGRGSPSEAMVGHLPFIPTKWLQDVFNWPLVI-----QMTDDKYLWKDL 136
b 6 DQRPILL-TGDRPTGQ-LHLGHFVGSLSRVLGQSHHQHLLADQAQALTQAD----- 57
Y 137 TLDQAVGDAVENAKDIIACGFDINKTFIF--SDLDYMGSSGFYKN---VVKIQKHVTFN 191
b 58 NPDKVRNILEVALDYLAVGIDPTKTTICVQSCLPALNLTWYLYNFVTVARLERNPTIK 117
Y 192 QVKGIFGFTSDCIGKISFPFAIQAPSPNSFPQIFRDRTDIQC-LIPCAIDQDPYFMT 250
b 118 SETQMRGFERDIPAGFLCYPAQAA-----DITAFKATVVPVGVGDQIPMIEQT 145
Y 251 RDVAPRIGYP-----KPAIHLSTFFPALQAGATMSASDPSIFLTDIAKQIKTV 301
b 166 NEIVRRVNRQIGDQLPECKALLSNMARLPQFDG-KAKMSKSLGN-TIVLNASDKOIKKA 223
Y 302 VNKHAFSGGRDITIEHRQFGNCDVDSVYMLTFFLEDDDKLEQIRKDYTSGLMTGELK 361
b 224 VNAMYTDPNHLRIEDDPQVGEV-----IVFTYLDADFDPNKEVVELKAHYRGGGIDGTVK 279
Y 362 KALIEVLQPLIAEHQARKKEVTDE 385
b 280 KRLEGVLKELITPIRERREBLAXD 303

```

RESULT 3

```

US-09-107-532A-5765
; Sequence 5765, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5765:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...348
; SEQUENCE DESCRIPTION: SEQ ID NO: 5765:
US-09-107-532A-5765

Query Match      8.8%; Score 186; DB 4; Length 348;
Best Local Similarity 24.9%; Pred. No. 4.3e-12;
Matches 82; Conservative 57; Mismatches 120; Indels 70; Gaps 14;
QY 88 LYTGGRSSSEAMVGHLPFIPTKWLQDVFNWPLVIQMTDDKYLWKDLTLDQAVGDAVE 147
DB 17 ILTGDRPTGK-LHLGHVGSLSKTR-----VAMQADESNQFLVMIADMQALTQNAK 65
QY 148 N-----AKDIIACGFDINKT--FISDLDYMGSSGFYKNVY---KIQKHVT-- 189
DB 66 NPEKVVSNVLQVALDYLAVGLDPAKSTLFIQSIQPELAELTWYLYNLVSVGRVNRPTVK 125
QY 190 --FNQVKGIFGFTSDCIGKISFPFAIQAPSPNSFPQIFRDRTDIQC-LIPCAIDQDPY 246
DB 126 TEIEQKK-----FGSVPTGFFIYFVSQAA-----DITAFKANLVPGVGDQKPM 169
QY 247 FRMTRD-----VAPRIGYPKAPLHSTFFPALQAGATMSASDPSNSIFLTD 293
DB 170 LEQTOEIVQSFNHTYGEVLVEPKGVFPKGMGR---LPGIDG-NGRKMSKSLGN-GIYISD 224
QY 234 TAKQIKTVNKHAFSGGRDITIEHRQFGNCDVDSVYMLTFFLEDDDKLEQIRKDYTSG 353
DB 225 PADVLQKKV-----MSMYTDPNHLHVQDPGQVGEVKNVFTYLDVFTDKNEA-LEEMKAHYRRG 280

```



QY 354 AMLTGELKALIEVLQPLIAEHQARRKEV 382  
 Db 281 GLGDVKIKRILIDVLEAFAPIRARREL 329

## RESULT 4

US-08-928-100-2  
 ; Sequence 2, Application US/08928100  
 ; Patent No. 6046174  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gentry, Danile  
 ; APPLICANT: Greenwood, Claire  
 ; APPLICANT: Lawlor, Elizabeth  
 ; TITLE OF INVENTION: NO. 6046174el trps  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,100  
 ; FILING DATE: 12-SEP-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 9619072.3  
 ; FILING DATE: 12-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gimmi, Edward R.  
 ; REGISTRATION NUMBER: 38,891  
 ; REFERENCE/DOCKET NUMBER: P31624-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-4478  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 341 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-928-100-2

Query Match 8.8%; Score 185.5; DB 3; Length 341;  
 Best Local Similarity 24.0%; Pred. No. 4.8e-12;  
 Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVNVPLVIQMTDDKEY-LWKDLTLDOAY 142  
 Db 3 KPIIL-TGDRPTGK-LHGHYVYGLSKNR-----VLLQEEKDYDMFVFLAQOQAL 49  
 QY 143 GDAVEN-----AKDIIACGFDINKTIF--SOLDYMGSSGFYKNVY---KIOK 186  
 Db 50 TDHAKDPQTIVESIGNVALDYLAAGLDPNKSTIFIQSQIPELAELSMYWNVLARLER 109  
 QY 187 HVTFNQVKGIFGFTSDCIGKISFPAIQAA--PSFSNSFPQIFRDRTOICLIPCAIDQD 244  
 Db 110 NPTVKTETISQKGFSGISPTGELVYPIAQAADITAFKANY-----VPGVTDQK 156  
 QY 245 PYFRMTRD-----VAPRIGYKPEALLHSTFFPALQGAQTQMSASDPNSSIF 290  
 Db 157 PMIQTRTEIVRSFNAYNCDVLVEPEGIYPENE--RAGRLPLDGLG-NAKMSKS-LNNGIY 212  
 QY 291 LTDATKQIKTKVKNHAFSGGDRDTIEHQFGGNCVDVYSFMYLTFF--LEDDKLEQIKR 348

Db 213 LAUDADTLRKKVMSMYTDPDHLRVEDPGKIGN-----MVFHLDVFGRPEDAQEIADME 268  
 QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARRKEVTEI 386  
 Db 269 FYQGGGLGDVTKRYLLEILERELGPIRERRIEFAKDM 306

## RESULT 5

US-09-492-581-2  
 ; Sequence 2, Application US/09492581  
 ; Patent No. 6346409  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gentry, Danile  
 ; APPLICANT: Greenwood, Claire  
 ; APPLICANT: Lawlor, Elizabeth  
 ; TITLE OF INVENTION: NO. 6346409el trps  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/492,581  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,100  
 ; FILING DATE: 12-SEP-1997  
 ; APPLICATION NUMBER: 9619072.3  
 ; FILING DATE: 12-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gimmi, Edward R.  
 ; REGISTRATION NUMBER: 38,891  
 ; REFERENCE/DOCKET NUMBER: P31624-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-4478  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 341 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-492-581-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
 Best Local Similarity 24.0%; Pred. No. 4.8e-12;  
 Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLQDVNVPLVIQMTDDKEY-LWKDLTLDOAY 142  
 Db 3 KPIIL-TGDRPTGK-LHGHYVYGLSKNR-----VLLQEEKDYDMFVFLAQOQAL 49  
 QY 143 GDAVEN-----AKDIIACGFDINKTIF--SOLDYMGSSGFYKNVY---KIOK 186  
 Db 50 TDHAKDPQTIVESIGNVALDYLAAGLDPNKSTIFIQSQIPELAELSMYWNVLARLER 109  
 QY 187 HVTFNQVKGIFGFTSDCIGKISFPAIQAA--PSFSNSFPQIFRDRTOICLIPCAIDQD 244  
 Db 110 NPTVKTETISQKGFSGISPTGELVYPIAQAADITAFKANY-----VPGVTDQK 156  
 QY 245 PYFRMTRD-----VAPRIGYKPEALLHSTFFPALQGAQTQMSASDPNSSIF 290

Db 157 PMIEQREIVRSFNNAVCNCDVLEPGIYENB--RAGRLPGLDG-NAKMSKS-LNNGIY 212  
QY 291 LDTAKQIKTKVNHAFSGGRDTIEHRQFGNCDVDSFMYLTFF--LEDDDKLEQIRK 348  
Db 213 LADDAJTLRKVMKSMYTDPDHIEVEDPGKIEGN---MVPHYLDVFGRPDAQEIADMKE 268  
QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARRKEVTDEI 386  
Db 269 RYORGGLGDVKKRYLLEILRELGPERRERIEFAKDM 306

## RESULT 6

US-09-425-666-2

; Sequence 2, Application US/09425666

; Patent No. 6416976

; GENERAL INFORMATION:

; APPLICANT: Gentry, Danile

; APPLICANT: Greenwood, Claire

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: No. 6416976el trps

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESS: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/425,666

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/928,100

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Gimmi, Edward R

; REGISTRATION NUMBER: 38,891

; REFERENCE/DOCKET NUMBER: P31624-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-4478

; TELEFAX: 610-270-5090

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 341 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-425-666-2

Query Match 8.8%; Score 185.5; DB 4; Length 341;  
Best Local Similarity 24.0%; Pred. No. 4,8e-12;  
Matches 81; Conservative 60; Mismatches 128; Indels 69; Gaps 15;

QY 84 KPPLYTGRGPSSEAMHVLHPIFTKMLQDVFNPLVIQMTDDKY-LWKDLTLDQAY 142  
Db 3 KPILL-TGRPTGK-LHGHVYSLNR-----VILQEEKYDMFVFLADQAL 49  
QY 143 GRAVEN-----AKDIIACGFDINKTFIF--SLDLYMGSSGFGYKNVY---KIQK 186  
Db 50 TDHAKDPQTVESIGNVALDVLAVGLDENKSTFIQSQIPELAEIWMYNNVLSLARLER 109  
QY 187 HVTFNQVKGIFGTDSCCKGKSFPAIAQA--PSFNSPFIQFRDRTDIQCLIPCAIDQD 244  
Db 110 NPTVKTETISQKFGSGSIPTGFLVPIAQAADITAFKANY-----VPGVGTDX 156

QY 245 PYFRMTRD-----VAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIF 290  
Db 157 PMIEQREIVRSFNNAVCNCDVLEPGIYENB--RAGRLPGLDG-NAKMSKS-LNNGIY 212  
QY 291 LDTAKQIKTKVNHAFSGGRDTIEHRQFGNCDVDSFMYLTFF--LEDDDKLEQIRK 348  
Db 213 LADDAJTLRKVMKSMYTDPDHIEVEDPGKIEGN---MVPHYLDVFGRPDAQEIADMKE 268  
QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARRKEVTDEI 386  
Db 269 RYORGGLGDVKKRYLLEILRELGPERRERIEFAKDM 306

## RESULT 7

US-09-134-000C-5784

; Sequence 5784, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: LYNE Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; FILE REFERENCE: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,000C

; PRIOR FILING DATE: 1998-08-13

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 5784

; LENGTH: 335

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

; US-09-134-000C-5784

Query Match 8.2%; Score 174; DB 4; Length 335;  
Best Local Similarity 23.5%; Pred. No. 9,1e-11;  
Matches 81; Conservative 59; Mismatches 123; Indels 82; Gaps 16;

QY 88 LYTGEPSSEAMHVGHLIP-----FIFTKMLQDVFNPLVIQMTDDKY 131  
Db 5 ILTGDRPTGK-LHGHVYSLNRKXVEMQADPTMQLFVMIADLQ-----ALTDNAR- 53  
QY 132 LWKDLTLDQAYGDAVENAKDIIACGFDINKTFIF--SLDLYMGSSGFGYKNVY---KIQK 186  
Db 54 -----NPEKYSANVLEVALDVLAVGLDPTKTFIIOQIPQLAELTMVYLNLTTSVR 108  
QY 187 HVT-----FNQVKGIFGTDSCCKGKSFPAIAQAAPSPNSPFIQFRDRTDIQCLIPCAI 241  
Db 109 NPTVKAIEQKK-----FGEVPTGFFIYVVSQA-----DITAFQANLVVPG 152  
QY 242 DQDPYFRMTRD-----VAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSS 288  
Db 153 DQKPMLEQAQEI VHSFNQTYGEVLRPEAVLPKMGCR---LPGIDG-NGRMKSXLGN-G 207  
QY 289 IFLTDTAQIKTKVNHAFSGGRDTIEHRQFGNCDVDSFMYLTFFLEDDDKLEQIRK 348  
Db 208 IYLSDPRAEVVQKKVMSMYTDPNHIEVEDPGQVEGN---MVFYLDVFGKKNKEYIELKE 263  
QY 349 DYTSGAMLTGELKALIEVLQPLIAEHQARRKEV--TDEIVKFEFM 391  
Db 264 HYRHGGLGDVKKRYLIDVLEELAPIRRERRELAKNPEAIMEML 308

## RESULT 8

US-08-743-130A-39

; Sequence 39, Application US/08743130A

; Patent No. 5871987

; GENERAL INFORMATION:

; APPLICANT: Sasanfar, Mandana

; APPLICANT: Gallant, Paul L.

; APPLICANT: Shen, Xiayu

; APPLICANT: Tao, Nianjun

APPLICANT: Tao, Jianshi  
 APPLICANT: Houman, Fariba  
 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/743,130A  
 FILING DATE: 01-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook Esq., David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-12  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 861-6240  
 TELEFAX: (617) 861-9540  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 409 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: linear  
 TOPOLOGY: linear  
 US-08-743-130A-39

Query Match 7.8%; Score 164.5; DB 2; Length 409;  
 Best Local Similarity 20.7%; Pred. No. 1.5e-09;  
 Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;  
 QY 76 VLDAYENK-KPFVLYTGRGSSSEAMHVGHLIPFI-----FTKWLDQV-----F 117  
 DB 27 IKDVLKERNRPVKIYWGTAPTGPK-HCGYFVPMKLAHFLKAGEVTVLLADLHAFLDNM 85  
 QY 118 NVPLVIQMTDDEKYLKDLTDQAYGDAVENAKDIACGDFINKTFESLDLYNGMSSGF 177  
 DB 86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGDYV---MDLFKLSNIV 142  
 QY 178 YKNVVK-----IQKHVTFNQVKGIFGFTDSDCIGKISFPALQAAPSFNSFPQIFDRD 232  
 DB 143 SQNDAKRAGADVVKQVANPLLSGLI-----YPLMQA-----IDEEHLG 180  
 QY 233 IQCLIPCAIDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLT 292  
 DB 181 VDAQFG-GVQDKIFVLAEENLPSIGYKRAHLMNPMVPSGL-GQGGKMSASDPNSKIDII 238  
 QY 293 DTAKOIKTKVKNHAFSGG--RDT-----IEHRQFGGN 323  
 DB 239 EEPKVVKKKNSAYCAPGELKDNGLIAFYVIQPIAELKTGVEGAFKLDIDRPEKYGG- 297  
 QY 324 CDVDSFMYLTFLEDDDDKLEQIRKDYTSGLMTGLKXALI-----EVLQPLIAEHQARR 379  
 DB 298 ---PLSY-----DSIEQLKADFVQDKLAPDLKSGVADKINELLAPTRAEPFESS- 343  
 QY 382 KEVTEIVKEMTPRK 395  
 DB 344 -----EEFQVQAK 351

RESULT 9

US-08-743-130A-2

; Sequence 2, Application US/08743130A

; Patent No. 5871987

GENERAL INFORMATION:  
 APPLICANT: Saseanfar, Mandana  
 APPLICANT: Gallant, Paul L.  
 APPLICANT: Shen, Xiaoyu  
 APPLICANT: Tao, Nianjun  
 APPLICANT: Tao, Jianshi  
 APPLICANT: Houman, Fariba  
 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02173  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/743,130A  
 FILING DATE: 01-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brook Esq., David E.  
 REGISTRATION NUMBER: 22,592  
 REFERENCE/DOCKET NUMBER: CPI95-12  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 861-6240  
 TELEFAX: (617) 861-9540  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 409 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: linear  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-743-130A-2

Query Match 7.7%; Score 162.5; DB 2; Length 409;  
 Best Local Similarity 20.7%; Pred. No. 2.5e-09;  
 Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;  
 QY 76 VLDAYENK-KPFVLYTGRGSSSEAMHVGHLIPFI-----FTKWLDQV-----F 117  
 DB 27 IKDVLKERNRPVKIYWGTAPTGPK-HCGYFVPMKLAHFLKAGEVTVLLADLHAFLDNM 85  
 QY 118 NVPLVIQMTDDEKYLKDLTDQAYGDAVENAKDIACGDFINKTFESLDLYNGMSSGF 177  
 DB 86 KAPLEVVKYRAKYEFVVKAILKSNVPIERLKFVVGSSYQKGDYV---MDLFKLSNIV 142  
 QY 178 YKNVVK-----IQKHVTFNQVKGIFGFTDSDCIGKISFPALQAAPSFNSFPQIFDRD 232  
 DB 143 SQNDAKRAGADVVKQVANPLLSGLI-----YPLMQA-----IDEEHLG 180  
 QY 233 IQCLIPCAIDQDPYFRMTDVAIRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLT 292  
 DB 181 VDAQFG-GVQDKIFVLAEENLPSIGYKRAHLMNPMVPSGL-GQGGKMSASDPNSKIDII 238  
 QY 293 DTAKOIKTKVKNHAFSGG--RDT-----IEHRQFGGN 323  
 DB 239 EEPKVVKKKNSAYCAPGELKDNGLIAFYVIQPIAELKTGVEGAFKLDIDRPEKYGG- 297  
 QY 324 CDVDSFMYLTFLEDDDDKLEQIRKDYTSGLMTGLKXALI-----EVLQPLIAEHQARR 379  
 DB 298 ---PLSY-----DSIEQLKADFVQDKLAPDLKSGVADKINELLAPTRAEPFESS- 343  
 QY 380 KEVTEIVKEMTPRK 395  
 DB 344 -----EEFQVQAK 351

```

RESULT 10
US-09-489-039A-8660
; Sequence 8660, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.3004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8660
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8660

Query Match      7.4%; Score 157.5; DB 4; Length 385;
Best Local Similarity 24.4%; Pred. No. 8.3e-09;
Matches 87; Conservative 53; Mismatches 142; Indels 75; Gaps 17;

QY 60 HFLRR---GIFFSHRDMQVLDAYENKX-----PFYLTGGRSPSEAHVGHLPPIFT 110
DB 25 HFQSRKALVFSR-----LSAQSDKESLMNHPTILTDGRTGQ-LHLGHVYVGLRQ 78
QY 111 KWLQDVNVPLVIQMTDDEKYLWKDLTLQDAYGDAVENAK-----DIACGFDI 159
DB 79 R-----VALQHNHQFILIADL---QGLTNGSNPKISHHILEVMADYLAUGLDT 126
QY 160 NKTIFP--SDLDYMGSSGFYKNVY---KIQKHVTFNOVKIGFOTSDCIGKISFPAIQ 214
DB 127 RLTTICQSLPALAEALSALYNNITVARVERNPTKNEIAQKGFARSLPVGFLAYPISQ 185
QY 215 AAPFSFNSFPQIFDRDRTDQC-LIPCALDQDPYFRTMDVA-----PRIGYPKAL 264
DB 187 AA-----DITAFKALVPVGDQLEPMEIQNEIVHKNSLTGEPVLSHCKALL 234
QY 265 LHSFTFPALQAGTQKMSADPNSIFLDTAKQIKTKVKNHAFSGGRDTEHHRQFG--G 322
DB 235 SEVSRLEPGVDG-NAKMSKSLGN-TLTSATEEEI-----HNAVY-AMYTPTHLRVSDPG 286
QY 323 NCDVDVSEMYLTLEFDDDKLEQIRKDYTSGLMTGELKALIEVLQPLIAEHQARR 379
DB 287 HVEGNVVFYLDFAHSDKARVAEMKTHYQRGGLGDRQCKNELETCLQTLLAPIRER 343

RESULT 11
US-09-198-452A-857
; Sequence 857, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE REFERENCE: thereof and uses thereof, in particular for the diagnosis, prevention
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 857
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-857

Query Match      7.2%; Score 152; DB 4; Length 344;
Best Local Similarity 23.1%; Pred. No. 2.8e-08;
Matches 79; Conservative 50; Mismatches 127; Indels 86; Gaps 16;

```

```

QY 82 NKKPFVLYTGRGSPSEAHVGHLPPIFTKWLQDVNVPLVIQMTDEKY----- 131
DB 2 NKKKVJ--TGDRTGK-LHLGH-----WVGSIKN-----RLELQNSPEYDCFFILADLH 48
QY 132 -----LWKOLTLQDAYGDAVEN-----AKDIIACGFDINKFTIFSDLDYMGSSGFYKNV 181
DB 49 TLTTIKRKEVLD-----VDNHIYEVADLWSLVGIDFTKSIY-----LQSAPEIYELH 98
QY 182 VKIQKHVTFNOVKIGFOTD-----SDCIGKISFPAIQAPFSFNSFPQIFDRDT 231
DB 99 LFPMSLISINRYMGIPSLKDMARNASIEGSLSYGLIGYPILOSA----- 143
QY 232 DIQC-----LIPCALDQDPYFRTMDVA---YKPAALLHS---TEFPALQAGTQK 279
DB 144 DILLAKAQFVPVGNKNEAHVELTRDARNFNLYQGVFPEPEVLQCEJ--SLVGIDG--QKG 202
QY 280 MSADPNSIFLDTAKQIKTKVKNHAFSGGRDTEHHRQFGNCDCVDVSEMYLTLEFED 339
DB 203 MSKS-ANNAIYLSDSDATITEKVRKMYTDPENRIRATTPGVEGN-----PLFIYHDIPEH 257
QY 340 DKLQIRKDYTSGLMTGELKALIEVLQPLIAEHQARRKE 381
DB 258 KDEVEEFKARYQGCIOIKIEVKARLAEEELIHLKPIKERRSE 299

RESULT 12
US-08-705-868-4
; Sequence 4, Application US/08705868
; Patent No. 5885798
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Ru-Young, Janice
; APPLICANT: Murry, Synn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,868
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1184599
US-08-705-868-4

Query Match      6.3%; Score 132.5; DB 2; Length 388;
Best Local Similarity 20.8%; Pred. No. 5.4e-06;

```



Db 129 ERNPTVKHILQK-----NLSRSLPAGELTPVVSQAADITAFSAD-----IV 170  
Qy 238 PCAIDQDPYFRMTRDVAPR-----IGYP-----KPALLHSTFPFALQGAQTKMSASDPNSS 288  
Db 171 PAGEDQLPMIBQTNIEIVTKINSLIQCPVLTSCVKVVGVQGRPLPGTDGS-GRMSKSLGN-T 228  
Qy 289 IFLTDTAKQIKTKVKNHAFSGGRDTEBHRQFGGNCVDVDSFMYLTFLEDDDKLEQIRK 348  
Db 229 INLSSTADEIK-----KAVYSMTDPQHIDVASPGHIEGNVFTYLDACQDKATVTAMKA 284  
Qy 349 DYTSGAMLTGELK-KALI-----EVLQPL-----IAEQARRKEVTDE 385  
Db 285 HYQGGCL--GDMCKKAMLDILOELLQPIREKRAQLINDKAYLLQVIXEGSGDKAKEVTQQ 342  
Qy 386 IVKE 389  
Db 343 KLDE 346

RESULT 15  
US-09-543-681A-6769  
; Sequence 6769, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543.681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 6769  
; LENGTH: 349  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-6769

Query Match 5.6%; Score 118; DB 4; Length 349;  
Best Local Similarity 21.0%; Pred. No. 0.0002;  
Matches 80; Conservative 46; Mismatches 107; Indels 148; Gaps 20;  
Qy 83 KK3EVLTYTGRGSSSEAMVGHLPFTKWLQDVFNVLVIQMTDEKYLW-----KDLT 137  
Dt 17 QKPI-VFSGAQESGE-LTIGNYMGAL-ROW-----VQMDDYDCIYCIVDQHAIT 63  
Qy 138 LDQAYGDVAVENAKD-----IIACGFINDKTFIFSDLDYMGSSGFYKVVYKIQKHVT----- 189  
Dt 64 VQDPTFLKRTLDLTALYIACIDPEKSTIF-----VQSHVPQHAQ 105  
Qy 190 -----FNQVKGIFGFTSDC-----IGKISFPFQIAQAPSFNSFPQIFRDRTD 232  
Dt 106 LGWALNCVTFYFGLSRMTQFKDKSARHAENINAGLPDYPVLMAA-----DILLYQTN 157  
Qy 233 LQCLIPCALDQDPYFRMTRDVAPRIG--YKPKALLHSTFPF-----ALQGAQTKMSAS 283  
Dt 158 ---QPVGIDQKHLELSGSDIAORFNAYIGDIFTVPDPPIPKGARVMALQDPFAKKMS 214  
Qy 284 DPNSS---IFLTD---TAKQIKTKV-----NKEAFSG-----GRDTI 314  
Db 215 DNNRNVIALLDPPKAAKKIKRAVTDSEPPRVAYDLENKAGVSNLLDILAGVTGTKTP 274  
Qy 315 BEHROFGNCVDVDSFMYLTFLEDDDKLEQIRKDYTGAMLTGELKKALIEVLQPLIAE 374  
Db 275 ELEAEFECK-----MY-----GHLKGAFAVAVSDMLTN 302  
Qy 375 HQAR-----RKEVTDEIVKE 389  
Db 303 IQERFNTFRNDEALINKIMKE 323

Search completed: August 24, 2004, 18:04:24

Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 17:18:04 ; Search time 126 Seconds  
(without alignments)  
899.219 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471  
Perfect score: 2116  
Sequence: 1 SNHGPDATBAEDFVDPWTV.....VTDEIVKFEFTPRKLSDFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: Geneseq\_29Jan04:\*
  - 2: Geneseq\_1980s:\*
  - 3: Geneseq\_1990s:\*
  - 4: Geneseq\_2000s:\*
  - 5: Geneseq\_2001s:\*
  - 6: Geneseq\_2002s:\*
  - 7: Geneseq\_2003s:\*
  - 8: Geneseq\_2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2116   | 100.0       | 401    | 5     | AAG79549 TrpRS T1  |
| 2          | 2116   | 100.0       | 415    | 4     | AAB47617 Human sup |
| 3          | 2116   | 100.0       | 415    | 5     | AAB47617 Human sup |
| 4          | 2116   | 100.0       | 415    | 5     | AAG79548 His6-tag  |
| 5          | 2116   | 100.0       | 415    | 6     | ABU72386 Human sup |
| 6          | 2116   | 100.0       | 437    | 4     | AAB47616 Human min |
| 7          | 2116   | 100.0       | 437    | 5     | AAB47616 Human min |
| 8          | 2116   | 100.0       | 437    | 5     | AAG79547 His6-tag  |
| 9          | 2116   | 100.0       | 437    | 6     | ABU72385 Human min |
| 10         | 2116   | 100.0       | 484    | 4     | AAB47615 Human ful |
| 11         | 2116   | 100.0       | 484    | 5     | AAB47615 Human ful |
| 12         | 2116   | 100.0       | 484    | 5     | AAG79546 Full leng |
| 13         | 2116   | 100.0       | 484    | 6     | ABU72384 Human ful |
| 14         | 2101   | 99.3        | 471    | 7     | ADE25762 Human pro |
| 15         | 2101   | 99.3        | 471    | 8     | ADE76998 Human pro |
| 16         | 2101   | 99.3        | 475    | 3     | AAB58220 Lung carc |
| 17         | 2096   | 99.1        | 471    | 2     | AAY05372 Human HCM |
| 18         | 1388   | 94.3        | 378    | 5     | AAG79541 TrpRS T2  |
| 19         | 1388   | 94.3        | 392    | 4     | AAB47618 Human ica |
| 20         | 1388   | 94.0        | 392    | 5     | AAB47618 Human ica |
| 21         | 1388   | 94.0        | 392    | 5     | AAG79544 His6-tag  |
| 22         | 1388   | 94.0        | 392    | 6     | ABU72387 Human min |
| 23         | 1973   | 93.2        | 378    | 6     | ABU72387 Human min |
| 24         | 1938   | 91.6        | 475    | 7     | ADB79825 Mouse put |
| 25         | 1370.5 | 64.8        | 430    | 4     | ABB64621 Drosophil |

|    |        |      |     |   |          |
|----|--------|------|-----|---|----------|
| 26 | 1370.5 | 64.8 | 430 | 4 | ABB67203 |
| 27 | 1304   | 61.6 | 402 | 3 | AAG23698 |
| 28 | 1304   | 61.6 | 426 | 3 | AAG23697 |
| 29 | 1218.5 | 57.6 | 424 | 4 | AAB66931 |
| 30 | 1200.5 | 56.7 | 424 | 5 | ABP73795 |
| 31 | 1163   | 55.0 | 432 | 6 | ABR53898 |
| 32 | 1125   | 53.2 | 433 | 6 | ABJ26487 |
| 33 | 1107.5 | 52.3 | 456 | 7 | ADB70160 |
| 34 | 968    | 45.7 | 292 | 3 | AAG23699 |
| 35 | 831    | 39.3 | 173 | 3 | AAB58517 |
| 36 | 803    | 37.9 | 385 | 4 | AAB96409 |
| 37 | 414.5  | 19.6 | 179 | 6 | A325887  |
| 38 | 328    | 15.5 | 85  | 5 | AAE13515 |
| 39 | 328    | 15.5 | 85  | 6 | ABU72416 |
| 40 | 292    | 13.8 | 85  | 6 | ABU72417 |
| 41 | 273.5  | 12.9 | 85  | 6 | ABU72419 |
| 42 | 263    | 12.4 | 85  | 6 | ABU72413 |
| 43 | 226.5  | 10.7 | 142 | 3 | AAB58219 |
| 44 | 203.5  | 9.6  | 341 | 5 | A3P26964 |
| 45 | 195    | 9.2  | 337 | 6 | ABU17282 |

ALIGNMENTS

RESULT 1  
AAG79549  
ID AAG79549 standard; protein; 401 AA.  
XX  
AC AAG79549;  
XX  
DT 10-DEC-2002 (first entry)  
XX  
DE TrpRS T1 polypeptide.  
XX

T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation; neovascular eye disease; age-related macular degeneration; ocular complication; diabetes; rubecotic glaucoma; retinopathy; pterygia; keratitis; ischaemic retinopathy; sickle cell; pathologic myopia; ocular histoplasmosis; pterygia; T1; punctate innerchoriopathy; retinal degeneration; growth factor; vascularisation; vascular endothelial cell function; angiogenesis.

Homo sapiens.

WO200267970-A1.

06-SEP-2002.

22-FEB-2002; 2002WO-US005185.

23-FEB-2001; 2001US-0270951P.

(SRI ) SCRIPPS RES INST.

Schimmel P, Wakasugi K, Friedlander M;

WPI; 2002-698635/75.

New polypeptides derived from human tryptophanyl-tRNA synthase, useful for inhibiting ocular neovascularization in a patient, or for treating neovascular eye diseases, e.g. rubecotic glaucoma, retinopathy, keratitis, or pterygia.

Example 1; Page 78-79; 83pp; English.

This sequence represents a novel cleavage product, T1, of recombinant human tryptophanyl-tRNA synthase (TrpRS). A related cleavage product, T2, is water soluble and comprises residues 94-471 of full length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting ocular neovascularisation in a patient. The T2 polypeptide is useful for treating neovascular eye diseases, e.g. age-related macular degeneration, ocular complications of diabetes, rubecotic glaucoma, retinopathy of



CC prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell),  
 CC pathological myopia, ocular histoplasmosis, pterygia, or punitate  
 CC innerchoroidopathy. This polypeptide is particularly useful for treating  
 CC retinal degeneration to prevent the damaging effects of trophic and  
 CC growth factors, and for promoting vascularisation to retard retinal  
 CC degeneration by enhancing blood flow to cells. These are also useful for  
 CC regulating vascular endothelial cell function, and in particular, for  
 CC inhibiting angiogenesis  
 XX  
 XX Sequence 401 AA;

Query Match 100.0%; Score 2116; DB 5; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 Db 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 QY 61 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 120  
 Db 61 FLRGIFFSHRDMQVLDAYENKPFYLTGRGSSSEAMHVGHLIPFTKWLQDVFNVP 120  
 QY 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 Db 121 LVIQMTDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 QY 181 VKIQRHVTNFQVKGIFGFTSDCIKISFPAIQAAAPSFNSFPQIFRDTDIQCLIPCA 240  
 Db 181 VKIQRHVTNFQVKGIFGFTSDCIKISFPAIQAAAPSFNSFPQIFRDTDIQCLIPCA 240  
 QY 241 IDQDPYFMTROVAPRIGYKPKALHSTFPALQGAQTMSASDPSNSIFLDTAKQIKT 300  
 Db 241 IDQDPYFMTROVAPRIGYKPKALHSTFPALQGAQTMSASDPSNSIFLDTAKQIKT 300  
 QY 301 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 Db 301 KVNKHAFGSGGRTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401  
 Db 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401

RESULT 2  
 AAB47617  
 D AAB47617 standard; protein; 415 AA.  
 X AAB47617;

07-JAN-2002 (first entry)  
 Human supermini TrpRS.

Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain;  
 vascular endothelial cell function; burn; plastic surgery; abdomen;  
 polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 dermal ulcer; diabetic ulcer; endothelialization;  
 tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

Homo sapiens.

WC200174841-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US008966.

31-MAR-2000; 2000US-0193471P.

(SCRI) SCRIPPS RES INST.

PI

Schimmel P, Wakasugi K;

XX

WPI; 2001-626377/72.

DR

N-PSDB; AAB43634.

XX

New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

PT

vascular endothelial function, in particular for regulating angiogenesis,

PT

tumor metastasis and treating myocardial infarction.

XX

Disclosure; Page 129-30; 150pp; English.

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

DE Human supermini tryptophanyl t-RNA synthetase in pET20B.

XX Human: tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;  
KW TrpRS; vascular endothelial cell function; angiogenesis; wound healing;  
KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
KW angiography; gene therapy; tumour; inflammation; vascular permeability;  
KW rheumatoid arthritis; psoriasis; diabetic retinopathy.

IX Homo sapiens.

IX WO2002175078-A1.

IX 11-OCT-2001.

IX 21-MAR-2001; 2001WO-US008975.

IX 31-MAR-2000; 2000US-0193471P.

IX (SCRI ) SCRIPPS RES INST.

IX Schimmel P, Wakasugi K;

IX WPI; 2002-010784/01.

IX N-PSDB; A022484.

IT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
IT regulating vascular endothelial cell function, preferably angiogenesis,  
IT is useful for treating solid tumor or suppressing tumor metastasis in  
IT mammal.

PS Example 1; Page 129-130; 149pp; English.

XX The patent discloses human aminoacyl tRNA synthetases, particularly  
CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
CC fold nucleotide binding domain and polynucleotides encoding them. The  
CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
CC sequences are useful for regulating vascular endothelial cell function,  
CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
CC healing agents for re-vascularising damaged tissues. They are useful for  
CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
CC sequences can also be used in plastic surgery when reconstruction is  
CC required following a burn, other trauma, or even for cosmetic purposes.  
CC Angiogenic TrpRS is also used in association with surgery and following  
CC the repair of cuts, for promoting endothelialisation in vascular graft  
CC surgery and for repairing the damage of myocardial infarction and in  
CC conjunction with coronary bypass surgery by stimulating the growth of  
CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
CC immunoassays to detect the presence of tumours. They are also useful for  
CC blocking endogenous angiogenic activity and retard the growth of solid  
CC tumours. These antibodies may also be used to treat inflammation caused  
CC by increased vascular permeability. Inhibiting the activity of TrpRS by  
CC antisense technology is useful for preventing further growth or even  
CC regress solid tumours, and for treating rheumatoid arthritis, psoriasis,  
CC diabetic retinopathy, all of which are characterised by abnormal  
CC angiogenesis. The present sequence is human truncated tryptophanyl t-RNA  
CC synthetase (supermini TrpRS; residues 71-471 of full-length TrpRS  
CC protein) in pET20B

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 5; Length 415;  
Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATEAEEDFVDPMTVTQSSAKGIDYDKLIVRFGSSKIDKELINRIERATGCRPH 60  
DB 1 SNHGPDATEAEEDFVDPMTVTQSSAKGIDYDKLIVRFGSSKIDKELINRIERATGCRPH 60  
CY 61 FLRRGIFPSHRDMMQVLDAYENKXFFLYLTGCRPSSEAMHVGHLPFFTKWLQDVFNVP 120

DB 62 FLRRGIFPSHRDMMQVLDAYENKXFFLYLTGCRPSSEAMHVGHLPFFTKWLQDVFNVP 121  
QY 121 LVIQMTDDEKYLWDLTLDOAYGDAVENAKLIACGFDINKTFFISDLDYMGSSGFYKN 180  
DB 122 LVIQMTDDEKYLWDLTLDOAYGDAVENAKLIACGFDINKTFFISDLDYMGSSGFYKN 181  
QY 181 VVKIQKHVTNFQVKGIFGFTSDDCIGKISFPAIQAAPSFNSFPOIFRDRDTIOCLIPCA 240  
DB 182 VVKIQKHVTNFQVKGIFGFTSDDCIGKISFPAIQAAPSFNSFPOIFRDRDTIOCLIPCA 241  
QY 241 IDQPYFRMTRDVAPRIGYPKPALHSTFFPALQCAQTMSASDPNSSIFLTDTAQIKT 300  
DB 242 IDQPYFRMTRDVAPRIGYPKPALHSTFFPALQCAQTMSASDPNSSIFLTDTAQIKT 301  
QY 301 KVKHAFSGGRTIEHRQFGNCDVDVSMYLFLEDDDKLQIRKDYTSGAMLTGEL 360  
DB 302 KVKHAFSGGRTIEHRQFGNCDVDVSMYLFLEDDDKLQIRKDYTSGAMLTGEL 361  
QY 361 KXALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
DB 362 KXALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 402

RESULT 4

AAG79548

ID AAG79548 standard; protein; 415 AA.

AC AAG79548;

XX 10-DEC-2002 (first entry)

DE His6-tagged TrpRS T1 polypeptide.

XX T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation;  
KW neovascular eye disease; age-related macular degeneration;  
KW ocular complication; diabetes; rubecotic glaucoma; retinopathy;  
KW prematurity; keratitis; ischaemic retinopathy; sickle cell;  
KW pathological myopia; ocular histoplasmosis; pterygia; T1;  
KW puminate innerchoroidopathy; retinal degeneration; growth factor;  
KW vascularisation; vascular endothelial cell function; angiogenesis.  
XX Homo sapiens.

XX WO200267970-A1.

XX 06-SEP-2002.

XX 22-FEB-2002; 2002WO-US005185.

XX 23-FEB-2001; 2001US-0270951P.

XX (SCRI ) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K, Friedlander M;

XX WPI; 2002-698635/75.

XX N-PSDB; ABA00330.

XX New polypeptides derived from human tryptophanyl-tRNA synthase, useful  
XX for inhibiting ocular neovascularization in a patient, or for treating  
XX neovascular eye diseases, e.g. rubecotic glaucoma, retinopathy, keratitis,  
XX or pterygia.

XX Example 1; Page 71-72; 83pp; English.

XX This sequence represents a His6-tagged cleavage product, T1, of  
CC recombinant human tryptophanyl-tRNA synthase (TrpRS). A related cleavage  
CC product, T2, is water soluble and comprises residues 94-471 of full  
CC length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting  
CC ocular neovascularisation in a patient. The T2 polypeptide is useful for  
CC treating neovascular eye diseases, e.g. age-related macular degeneration,  
CC ocular complications of diabetes, e.g. age-related glaucoma, retinopathy of  
CC prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell),

C: pathological myopic, ocular histoplasmosis, pterygia, or punitate  
 C: innerchoroidopathy. This polypeptide is particularly useful for treating  
 C: retinal degeneration to prevent the damaging effects of trophic and  
 C: growth factors, and for promoting vascularisation to retard retinal  
 C: degeneration by enhancing blood flow to cells. These are also useful for  
 C: regulating vascular endothelial cell function, and in particular, for  
 C: inhibiting angiogenesis  
 X: Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 5; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q' 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
 D' 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 61  
 Q' 61 FLRGIFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 120  
 D' 62 FLRGIFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 121  
 Q' 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 D' 122 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 Q' 181 VKIQKHVTNFQVKGIFGFTSDDCIGKISFPALQAGTQKMSADPNSIFLTDIAKQIKT 240  
 D' 182 VKIQKHVTNFQVKGIFGFTSDDCIGKISFPALQAGTQKMSADPNSIFLTDIAKQIKT 241  
 Q' 241 IDQDPYFMRTRDVAPRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 300  
 D' 242 IDQDPYFMRTRDVAPRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 301  
 Q' 301 KVNKHFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 D' 302 KVNKHFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
 Q' 361 KXALIEVLOPLIAHQARRKEVTDEIVKEFMTPRKLSDFDQ 401  
 D' 362 KXALIEVLOPLIAHQARRKEVTDEIVKEFMTPRKLSDFDQ 402

RESULT 5  
 ABU72386  
 ID ABU72386 standard; protein; 415 AA.

XX AC ABU72386;

DT DT 16-JUN-2003 (first entry)

XX XX Human supermini Tryptophanyl tRNA synthetase/His tag.

DE DE Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytostatic;

XX KW vulnery; Rossmann fold nucleotide binding domain; chemokine; EMAP II;  
 KW vascular endothelial cell; solid tumour; myocardial infarction; enzyme;  
 KW endothelial monocyte-activating polypeptide II; tumour metastasis;  
 KW wound healing; dermal ulcer; endothelialisation; vascular graft surgery;  
 KW abdominal wound; coronary bypass surgery; gene therapy.

XX OS Homo sapiens.

OS Synthetic.

XX US2002182666-A1.

XX 35-DEC-2002.

XX 21-MAR-2001; 2001US-00813718.

XX 21-MAR-2001; 2001US-00813718.

XX (SCH/ ) SCHIMMEL P.  
 XX (WAKA/ ) WAKASUGI K.

XX PI  
 XX XX  
 DR DR  
 DR N-PSDB; ACA64107.

PT New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 PT Rossmann fold nucleotide binding domain or having chemokine activity  
 PT useful for e.g. for regulating angiogenesis and for treating myocardial  
 PT infarction.

XX Example 1; Page 50-51; 91pp; English.

CC The invention relates to an isolated polypeptide comprising a truncated  
 CC tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rossmann  
 CC fold nucleotide binding domain or having chemokine activity. The isolated  
 CC polypeptide is capable of regulating vascular endothelial cell function.  
 CC TrpRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 CC -activating polypeptide II, a proinflammatory cytokine)-like domain and  
 CC is similar in sequence to TrpRS. Also included are a polynucleotide  
 CC encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 CC comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 CC containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 CC TrpRS antibody, producing truncations of TrpRS by treating it with a  
 CC protease, and suppressing angiogenesis, solid tumours or a condition that  
 CC would benefit from decreased angiogenesis in a mammal by administering to  
 CC the mammal the composition comprising the TrpRS truncated protein. The  
 CC isolated polypeptide is useful for the preparation of a pharmaceutical  
 CC composition for transdermal, transmucosal, enteral or parenteral  
 CC administration. The truncated tRNA synthetase polypeptide is useful for  
 CC research, diagnostic, prognostic and therapeutic applications. The tRNA  
 CC synthetase are useful for regulating vascular endothelial cell function,  
 CC particularly for regulating angiogenesis, for treating myocardial  
 CC infarction and solid tumour, and for suppressing tumour metastasis.  
 CC Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 CC agents or for treating full thickness wounds such as dermal ulcers, in  
 CC the promotion of endothelialisation in vascular graft surgery, in the  
 CC treatment of abdominal wounds where there is a high risk infection, in  
 CC conjunction with coronary bypass surgery by stimulating the growth of the  
 CC transplanted tissue, and in gene therapy. The present sequence is a Trp  
 CC tRNA synthetase protein (full length, truncated or mutant) with a His  
 CC affinity tag

XX Sequence 415 AA;

Query Match 100.0%; Score 2116; DB 5; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
 D' 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 61  
 QY 61 FLRGIFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 120  
 D' 62 FLRGIFFSHRDMNQVLDAYENKPPVLYTGRGSSSEAMVGHLLPFIPTKWLQDVFNVP 121  
 QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 D' 122 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 QY 181 VKIQKHVTNFQVKGIFGFTSDDCIGKISFPALQAGTQKMSADPNSIFLTDIAKQIKT 240  
 D' 182 VKIQKHVTNFQVKGIFGFTSDDCIGKISFPALQAGTQKMSADPNSIFLTDIAKQIKT 241  
 QY 241 IDQDPYFMRTRDVAPRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 300  
 D' 242 IDQDPYFMRTRDVAPRIGYKPKALLHSTFFPALQAGTQKMSADPNSIFLTDIAKQIKT 301  
 QY 301 KVNKHFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 D' 302 KVNKHFSGGRDTIEHRQFGNCVDVSMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361

QY 361 KKALIEVLQPLIAEHQARRKEVTEIVKEFWTPKLSFDFQ 401  
 362 KKALIEVLQPLIAEHQARRKEVTEIVKEFWTPKLSFDFQ 402  
 RESULT 6  
 ID AAB47616 standard; protein; 437 AA.  
 AC AAB47616;  
 XX  
 DT 07-JAN-2002 (first entry)  
 DE Human mini TrpRS.  
 KW Tyrosyl-tRNA synthetase; TyrRS; Rossmann fold nucleotide binding domain;  
 KW vascular endothelial cell function; burn; plastic surgery; wound healing;  
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
 OS Homo sapiens.  
 PN WO200174841-A1.  
 PD 11-OCT-2001.

21-MAR-2001; 2001WO-US008966.

31-MAR-2000; 2000US-0193471P.

(SCRI ) SCRIPPS RES INST.

Schimmel P, Wakasugi K;

WPI; 2001-626377/72.

N-PSDB; AAB43603.

New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 vascular endothelial function, in particular for regulating angiogenesis,  
 tumor metastasis and treating myocardial infarction.

Disclosure; Page 123-24; 150pp; English.

The sequences given in AAB47615-18 show full length and truncated  
 versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
 the invention comprises a Rossmann fold nucleotide binding domain, and is  
 capable of regulating vascular endothelial cell function. It is of  
 approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
 full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing  
 angiogenesis to a graft, treating myocardial infarction, solid tumor, and  
 a condition that would benefit from increased or decreased angiogenesis  
 in a mammal, in particular humans. It is also useful in diagnosis and as  
 a wound healing agent for treating wounds such as dermal ulcers, diabetic  
 ulcers, burns and injuries and in plastic surgery when reconstruction is  
 required following a burn or for cosmetic purposes. It is particularly  
 useful in the treatment of abdominal wounds where there is high risk of  
 infection. Truncated TrpRS promotes endothelialization in vascular graft  
 surgery and is used in conjunction with angiography to administer the  
 angiogenic tRNA synthetase polypeptides or polynucleotides directly to  
 the lumen and wall of the blood vessel

Sequence 437 AA;

Query Match

Best Local Similarity 100.0%; Score 2116; DB 4; Length 437;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPH 60

DB 24 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPH 83  
 QY 61 FLRGIFFSHRDMNOVLDAVENKKEFYLYTGRGSSSEAMHVCHLIPFTKWLQDVFNVP 120  
 DB 84 FLRGIFFSHRDMNOVLDAVENKKEFYLYTGRGSSSEAMHVCHLIPFTKWLQDVFNVP 143  
 QY 121 LVIQMTDDEKYLKDLTLDAQYDAVENAKDIACGFDINKTFFPSDLDYMGSSSEFYKN 180  
 DB 144 LVIQMTDDEKYLKDLTLDAQYDAVENAKDIACGFDINKTFFPSDLDYMGSSSEFYKN 203  
 QY 181 VKIKQKVTNVQKGIFFGFTDSDCIGKISFPALIOAAPSFSNSPFIQFRDRTDIOCLIPCA 240  
 DB 204 VKIKQKVTNVQKGIFFGFTDSDCIGKISFPALIOAAPSFSNSPFIQFRDRTDIOCLIPCA 263  
 QY 241 IDODPYFRMTRDVAPRIGYKPKALLHSTFEPALOGAQTQMSASDPNSIFLDTAKQIKT 300  
 DB 264 IDODPYFRMTRDVAPRIGYKPKALLHSTFEPALOGAQTQMSASDPNSIFLDTAKQIKT 323  
 QY 301 KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLIGEL 360  
 DB 324 KVNKHAFFSGGRDTIEHRQFGGNCVDVSPWYLTFFLEDDDKLEIQIRKDYTSGAMLIGEL 383  
 QY 361 KKALIEVLQPLIAEHQARRKEVTEIVKEFWTPKLSFDFQ 401  
 DB 384 KKALIEVLQPLIAEHQARRKEVTEIVKEFWTPKLSFDFQ 424

RESULT 7

AAE13492

ID AAE13492 standard; protein; 437 AA.

AC AAE13492;

XX

DT 12-FEB-2002 (first entry)

XX

DE Human mini tryptophanyl t-RNA synthetase in pET20B.

XX

KW Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;  
 KW TyrRS; vascular endothelial cell function; angiogenesis; wound healing;  
 KW re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury;  
 KW diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction;  
 KW angiography; gene therapy; tumour; inflammation; vascular permeability;  
 KW rheumatoid arthritis; psoriasis; diabetic retinopathy.

OS Homo sapiens.

PN WO200175078-A1.

XX 11-OCT-2001.

PD 21-MAR-2001; 2001WO-US008975.

PF 31-MAR-2000; 2000US-0193471P.

PA (SCRI ) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K;

WPI; 2002-010784/01.  
 N-PSDB; AAB22483.

XX

PT Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of  
 regulating vascular endothelial cell function, preferably angiogenesis,  
 is useful for treating solid tumor or suppressing tumor metastasis in  
 mammal.

PS Example 1; Page 123-124; 149pp; English.

CC The patent discloses human aminocyl tRNA synthetases, particularly  
 CC truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann  
 CC fold nucleotide binding domain and polynucleotides encoding them. The  
 CC invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS  
 CC sequences are useful for regulating vascular endothelial cell function,

CC preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound  
 CC healing agents for re-vascularising damaged tissues. They are useful for  
 CC treating full-thickness wounds (e.g. dermal ulcers, including pressure  
 CC sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS  
 CC sequences can also be used in plastic surgery when reconstruction is  
 CC required following a burn, other trauma, or even for cosmetic purposes.  
 CC Angiogenic TrpRS is also used in association with surgery and following  
 CC the repair of cuts, for promoting endothelialisation in vascular graft  
 CC surgery and for repairing the damage of myocardial infarction and in  
 CC conjunction with coronary bypass surgery by stimulating the growth of  
 CC transplanted tissue. TrpRS is also used in conjunction with angiography.  
 CC TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in  
 CC immunoassays to detect the presence of tumours. They are also useful for  
 CC blocking endogenous angiogenic activity and retard the growth of solid  
 CC tumours. These antibodies may also be used to treat inflammation caused  
 CC by increased vascular permeability. Inhibiting the activity of TrpRS by  
 CC antisense technology is useful for preventing further growth or even  
 CC regressing solid tumours, and for treating rheumatoid arthritis, psoriasis,  
 CC diabetic retinopathy, all of which are characterised by abnormal  
 CC angiogenesis. The present sequence is human truncated tryptophanyl t-RNA  
 CC synthetase (mini TrpRS; residues 48-471 of full-length TrpRS protein)  
 CC protein in pBT20B  
 XX  
 XX Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 5; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNHGPDATAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
 DB 24 SNHGPDATAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
 QY 61 FLRGGIFFSHRDMNQVLDAYENKPFYLTGRGSSSEAMHVGHLPFIPTKWLQDVNVP 120  
 DB 84 FLRGGIFFSHRDMNQVLDAYENKPFYLTGRGSSSEAMHVGHLPFIPTKWLQDVNVP 143  
 QY 121 LVQMTDDEKYLWKDLTLQAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 180  
 DB 144 LVQMTDDEKYLWKDLTLQAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 203  
 QY 181 VKIKQHVTFNCKGIFGFTSDCKIGKISPPALQAAAPSFNSPQIFRDTDIQCLIPCA 240  
 DB 204 VKIKQHVTFNCKGIFGFTSDCKIGKISPPALQAAAPSFNSPQIFRDTDIQCLIPCA 263  
 QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKT 300  
 DB 264 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKT 323  
 QY 301 KVNKHAFFSGGRTTIEHRQFGNCVDVDSFMVLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 324 KVNKHAFFSGGRTTIEHRQFGNCVDVDSFMVLTFFLEDDDKLEQIRKDYTSGAMLTGEL 383  
 QY 361 KKALIEVLQPLIAEHOARKEVTDIVKGFMTKPLSDFDQ 401  
 DB 384 KKALIEVLQPLIAEHOARKEVTDIVKGFMTKPLSDFDQ 424

RESULT 8  
 AG79547  
 AG79547 standard; protein; 437 AA.

AG79547;  
 10-DEC-2002 (first entry)

His6-tagged mini-TrpRS polypeptide.

T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation;  
 neovascular eye disease; age-related macular degeneration;  
 ocular complication; diabetes; rubecotic glaucoma; retinopathy;  
 prematurity; keratitis; ischaemic retinopathy; sickle cell;  
 pathological myopic; ocular histoplasmosis; pterygia;

XX puniate innerchoroidopathy; retinal degeneration; growth factor;  
 XX vascularisation; vascular endothelial cell function; angiogenesis.  
 OS Homo sapiens.  
 PN WO200267970-A1.  
 XX 06-SEP-2002.  
 PD 22-FEB-2002; 2002WO-US005185.  
 XX 23-FEB-2001; 2001US-0270951P.  
 XX (SRI ) SRIEPS RES INST.  
 XX Schimmel P, Wakasugi K, Friedlander M;  
 XX WPI; 2002-698635/75.  
 DR N-PSDB; ABA00329.

XX New polypeptides derived from human tryptophanyl-tRNA synthase, useful  
 PT for inhibiting ocular neovascularization in a patient, or for treating  
 PT neovascular eye diseases, e.g. rubecotic glaucoma, retinopathy, keratitis,  
 PT or pterygia.

XX Example 3; Fig 1; 83pp; English.

XX This sequence represents a His6-tagged cleavage product, mini-TrpRS, of  
 CC recombinant human tryptophanyl-tRNA synthase (TrpRS). A related cleavage  
 CC product, T2, is water soluble and comprises residues 94-471 of full  
 CC length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting  
 CC ocular neovascularisation in a patient. The T2 polypeptide is useful for  
 CC treating neovascular eye diseases, e.g. age-related macular degeneration,  
 CC prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell),  
 CC pathological myopic, ocular histoplasmosis, pterygia, or puniate  
 CC innerchoroidopathy. This polypeptide is particularly useful for treating  
 CC retinal degeneration to prevent the damaging effects of trophic and  
 CC growth factors, and for promoting vascularisation to retard retinal  
 CC degeneration by enhancing blood flow to cells. These are also useful for  
 CC regulating vascular endothelial cell function, and in particular, for  
 CC inhibiting angiogenesis

XX Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 5; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-209;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
 DB 24 SNHGPDATAEEDFVDPWVTQSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
 QY 61 FLRGGIFFSHRDMNQVLDAYENKPFYLTGRGSSSEAMHVGHLPFIPTKWLQDVNVP 120  
 DB 84 FLRGGIFFSHRDMNQVLDAYENKPFYLTGRGSSSEAMHVGHLPFIPTKWLQDVNVP 143  
 QY 121 LVQMTDDEKYLWKDLTLQAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 180  
 DB 144 LVQMTDDEKYLWKDLTLQAYGDAVENAKDIIACGFDINKTIFSDLDYMGSSGFYKN 203  
 QY 181 VKIKQHVTFNCKGIFGFTSDCKIGKISPPALQAAAPSFNSPQIFRDTDIQCLIPCA 240  
 DB 204 VKIKQHVTFNCKGIFGFTSDCKIGKISPPALQAAAPSFNSPQIFRDTDIQCLIPCA 263  
 QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKT 300  
 DB 264 IDQDPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSSIFLDTAKQIKT 323  
 QY 301 KVNKHAFFSGGRTTIEHRQFGNCVDVDSFMVLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 324 KVNKHAFFSGGRTTIEHRQFGNCVDVDSFMVLTFFLEDDDKLEQIRKDYTSGAMLTGEL 383

QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEPMTPRKLSFDFQ 401  
 Db 384 KKALIEVLQPLIAHQARRKEVTDEIVKEPMTPRKLSFDFQ 424

## RESULT 9

ABU72385  
 ID ABU72385 standard; protein; 437 AA.

XX AC ABU72385;

XX DT 16-JUN-2003 (first entry)

XX Human mini Tryptophanyl tRNA synthetase/His tag.

XX Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytostatic;  
 KW vulnary; Rossmann fold nucleotide binding domain; chemokine; EMAP II;  
 KW vascular endothelial cell; solid tumour; myocardial infarction; enzyme;  
 KW endothelial monocyte-activating polypeptide II; tumour metastasis;  
 KW wound healing; dermal ulcer; endothelialisation; vascular graft surgery;  
 KW abdominal wound; coronary bypass surgery; gene therapy.

XX OS Homo sapiens.  
 OS Synthetic.

XX US2002182666-A1.

XX 05-DEC-2002.

XX 21-MAR-2001; 2001US-00813718.

XX 21-MAR-2001; 2001US-00813718.

XX (SCH/) SCHIMMEL P.  
 XX (WAKA/) WAKASUGI K.

XX Schimmel P, Wakasugi K;

XX WPI; 2003-340974/32.

XX N-PSDB; ACA64106.

XX New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 Rossmann fold nucleotide binding domain or having chemokine activity  
 useful for e.g. for regulating angiogenesis and for treating myocardial  
 infarction.

XX Example 1; Page 46-47; 91pp; English.

XX The invention relates to an isolated polypeptide comprising a truncated  
 tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rossmann  
 fold nucleotide binding domain or having chemokine activity. The isolated  
 polypeptide is capable of regulating vascular endothelial cell function.  
 CC TyRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 CC -activating polypeptide II, a proinflammatory cytokine)-like domain and  
 CC is similar in sequence to TrpRS. Also included are a polynucleotide  
 CC encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 CC 5' and 3' deletions of the TrpRS polynucleotide, a recombinant vector  
 CC comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 CC containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 CC TrpRS antibody, producing truncations of TrpRS by treating it with a  
 CC protease, and suppressing angiogenesis, solid tumours or a condition that  
 CC would benefit from decreased angiogenesis in a mammal by administering to  
 CC the mammal the composition comprising the TrpRS truncated protein. The  
 CC isolated polypeptide is useful for the preparation of a pharmaceutical  
 CC composition for transdermal, transmucosal, enteral or parenteral  
 CC administration. The truncated tRNA synthetase polypeptide is useful for  
 CC research, diagnostic, prognostic and therapeutic applications. The tRNA  
 CC synthetase are useful for regulating vascular endothelial cell function,  
 CC particularly for regulating angiogenesis, for treating myocardial  
 CC infarction and solid tumour, and for suppressing tumour metastasis.  
 CC Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 CC agents or for treating full thickness wounds such as dermal ulcers, in  
 CC the promotion of endothelialisation in vascular graft surgery, in the

CC treatment of abdominal wounds where there is a high risk infection, in  
 CC conjunction with coronary bypass surgery by stimulating the growth of the  
 CC transplanted tissue, and in gene therapy. The present sequence is a Trp  
 CC tRNA synthetase protein (full length, truncated or mutant) with a His  
 CC affinity tag

XX SQ Sequence 437 AA;

Query Match 100.0%; Score 2116; DB 6; Length 437;

Best Local Similarity 100.0%; Pred. No. 1.2e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQREHH 60

Db 24 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFSSKIDKELINRIERATGQREHH 83

QY 61 FLERGIFFSHRDNMVLDAYENKPPFLYITGRGSSSEAMHVGHLIPFTKWLQOVFNVP 120

Db 84 FLERGIFFSHRDNMVLDAYENKPPFLYITGRGSSSEAMHVGHLIPFTKWLQOVFNVP 143

QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180

Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203

QY 181 VVKIQHVTNPKVKGIFGFTDSDCIGKISFPALQAPSPNSFPQIFRDRDTIOCLIPCA 240

Db 204 VVKIQHVTNPKVKGIFGFTDSDCIGKISFPALQAPSPNSFPQIFRDRDTIOCLIPCA 263

QY 241 IDQDPYFRMTDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLDTAKQIKT 300

Db 264 IDQDPYFRMTDVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLDTAKQIKT 323

QY 301 KVNKHAFFSGGRDTEIEHRQFGNCDVDVSMYLTFFLEDDKLEQIRKDYTSGLMTGEL 360

Db 324 KVNKHAFFSGGRDTEIEHRQFGNCDVDVSMYLTFFLEDDKLEQIRKDYTSGLMTGEL 383

QY 361 KKALIEVLQPLIAHQARRKEVTDEIVKEPMTPRKLSFDFQ 401

Db 384 KKALIEVLQPLIAHQARRKEVTDEIVKEPMTPRKLSFDFQ 424

## RESULT 10

AAB47615

ID AAB47615 standard; protein; 484 AA.

XX AC AAB47615;

XX DT 07-JAN-2002 (first entry)

XX Human full-length TrpRS.

XX Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;  
 KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.

XX Homo sapiens.

XX WO200174841-A1.

XX 11-OCT-2001.

XX 21-MAR-2001; 2001WC-US008966.

XX 31-MAR-2003; 2000US-0193471P.

XX (SCRI) SCRIPPS RES INST.

XX Schimmel P, Wakasugi K;

XX WPI; 2001-626377/72.

DR XX  
XX XX  
PT PT  
PT PT  
PT PT  
XX XX  
XX XX

N-PSDB; AAH43602.  
New human truncated tyrosyl-tRNA synthetase polypeptide for regulating vascular endothelial function, in particular for regulating angiogenesis, tumor metastasis and treating myocardial infarction.  
Disclosure; Page 117-19; 150pp; English.

The sequences given in AAB47615-18 show full length and truncated versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of the invention comprises a Rossmann fold nucleotide binding domain, and is capable of regulating vascular endothelial cell function. It is of approx. 40 kilo Dalton molecular weight and is produced by cleavage of full length TrpRS with polymorphonuclear leucocyte elastase. Truncated TrpRS is useful for regulating angiogenesis, tumor metastasis, enhancing angiogenesis to a graft, treating myocardial infarction, solid tumor, and a condition that would benefit from increased or decreased angiogenesis in a mammal, in particular humans. It is also useful in diagnosis and as a wound healing agent for treating wounds such as dermal ulcers, diabetic ulcers, burns and injuries and in plastic surgery when reconstruction is required following a burn or for cosmetic purposes. It is particularly useful in the treatment of abdominal wounds where there is high risk of infection. Truncated TrpRS promotes endothelialization in vascular graft surgery and is used in conjunction with angiography to administer the angiogenic tRNA synthetase polypeptides or polynucleotides directly to the lumen and wall of the blood vessel.

Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 4; Length 484;

Best Local Similarity 100.0%; Pred. No. 1.4e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
71 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
61 FLRRGIFFSHRDMQVLDAYENKPFYLYTGRGSPSEAMHVGHLIPFIFTKMLQDVFNVP 120  
131 FLRRGIFFSHRDMQVLDAYENKPFYLYTGRGSPSEAMHVGHLIPFIFTKMLQDVFNVP 190  
121 LVIQMTDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFFISDLDMGSSGFYKN 180  
131 LVIQMTDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFFISDLDMGSSGFYKN 250  
181 VVKIQHVTENQVKGIFGFTSDSDCKIGISFPATQAPSFNSFPQIFRDRDTQCLIPCA 240  
251 VVKIQHVTENQVKGIFGFTSDSDCKIGISFPALQAPSFNSFPQIFRDRDTQCLIPCA 310  
241 IDQDPYFMRTRDVAPRIGYPKALLHSTFPFALQAGATQMSADPNSSIFLTDTAKQIKT 300  
311 IDQDPYFMRTRDVAPRIGYPKALLHSTFPFALQAGATQMSADPNSSIFLTDTAKQIKT 370  
301 KVNKHAFGSGRDTIEHRQFGNCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
371 KVNKHAFGSGRDTIEHRQFGNCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430  
361 KKALIEVTLQPLIAEHQARRKEVTDIVKEFMTPRKLSPDFQ 401  
431 KKALIEVTLQPLIAEHQARRKEVTDIVKEFMTPRKLSPDFQ 471

RESULT 11

AAE13491

AAE13491 standard; protein; 484 AA.

AAE13491;

12-FEB-2002 (first entry)

Human tryptophanyl t-RNA synthetase (TrpRS) in pET20B.

Human; tryptophanyl-tRNA synthetase; TrpRS; tyrosyl t-RNA synthetase;

KW

KW

KW

KW

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

TyrRS; vascular endothelial cell function; angiogenesis; wound healing; re-vascularisation; dermal ulcer; pressure sore; venous ulcer; injury; diabetic ulcer; burn; plastic surgery; cosmetic; myocardial infarction; angiography; gene therapy; tumour; inflammation; vascular permeability; rheumatoid arthritis; psoriasis; diabetic retinopathy.

Homio sapiens.

WO200175078-A1.

11-OCT-2001.

21-MAR-2001; 2001WO-US008975.

31-MAR-2000; 2000US-0193471P.

(SCRI) SCRIPPS RES INST.

Schimmel P, Wakasugi K;

WPI; 2002-010784/01.

N-PSDB; AAE22482.

Novel truncated tryptophanyl-tRNA synthetase polypeptides capable of regulating vascular endothelial cell function, preferably angiogenesis, is useful for treating solid tumor or suppressing tumor metastasis in mammal.

Example 1; Page 117-119; 149pp; English.

The patent discloses human aminoacyl tRNA synthetases, particularly truncated tryptophanyl-tRNA synthetases (TrpRS) comprising a Rossmann fold nucleotide binding domain and polynucleotides encoding them. The invention also relates to tyrosyl t-RNA synthetases (TyrRS). TrpRS sequences are useful for regulating vascular endothelial cell function, preferably angiogenesis. Angiogenic TrpRS sequences are useful as wound healing agents for re-vascularising damaged tissues. They are useful for treating full-thickness wounds (e.g. dermal ulcers, including pressure sores, venous ulcers and diabetic ulcers), burns and injuries. TrpRS sequences can also be used in plastic surgery when reconstruction is required following a burn, other trauma, or even for cosmetic purposes. Angiogenic TrpRS is also used in association with surgery and following the repair of cuts, for promoting endothelialisation in vascular graft surgery and for repairing the damage of myocardial infarction and in conjunction with coronary bypass surgery by stimulating the growth of transplanted tissue. TrpRS is also used in conjunction with angiography. TrpRS DNAs are useful in gene therapy. TrpRS antibodies are used in immunoassays to detect the presence of tumours. They are also useful for blocking endogenous angiogenic activity and retard the growth of solid tumours. These antibodies may also be used to treat inflammation caused by increased vascular permeability. Inhibiting the activity of TrpRS by antisense technology is useful for preventing further growth or even regress solid tumours, and for treating rheumatoid arthritis, psoriasis, diabetic retinopathy, all of which are characterised by abnormal angiogenesis. The present sequence is human tryptophanyl t-RNA synthetase (TrpRS) in pET20B

Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 5; Length 484;

Best Local Similarity 100.0%; Pred. No. 1.4e-209;

Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60

71 SNHGPDTEAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130

61 FLRRGIFFSHRDMQVLDAYENKPFYLYTGRGSPSEAMHVGHLIPFIFTKMLQDVFNVP 120

131 FLRRGIFFSHRDMQVLDAYENKPFYLYTGRGSPSEAMHVGHLIPFIFTKMLQDVFNVP 190

121 LVIQMTDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFFISDLDMGSSGFYKN 180

131 LVIQMTDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFFISDLDMGSSGFYKN 250

181 VVKIQHVTENQVKGIFGFTSDSDCKIGISFPATQAPSFNSFPQIFRDRDTQCLIPCA 240

251 VVKIQHVTENQVKGIFGFTSDSDCKIGISFPALQAPSFNSFPQIFRDRDTQCLIPCA 310

241 IDQDPYFMRTRDVAPRIGYPKALLHSTFPFALQAGATQMSADPNSSIFLTDTAKQIKT 300

311 IDQDPYFMRTRDVAPRIGYPKALLHSTFPFALQAGATQMSADPNSSIFLTDTAKQIKT 370

301 KVNKHAFGSGRDTIEHRQFGNCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360

371 KVNKHAFGSGRDTIEHRQFGNCDVDSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430

361 KKALIEVTLQPLIAEHQARRKEVTDIVKEFMTPRKLSPDFQ 401

431 KKALIEVTLQPLIAEHQARRKEVTDIVKEFMTPRKLSPDFQ 471



Db 191 LVQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKN 250

QY 181 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPPAIQAAPSFNSPQIFRDRDTIOCLIPCA 240

Db 251 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPPAIQAAPSFNSPQIFRDRDTIOCLIPCA 310

QY 241 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 300

Db 311 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 370

QY 301 KVKHAFSGGRDTIEHRQFGNCVDVDSFWYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360

Db 371 KVKHAFSGGRDTIEHRQFGNCVDVDSFWYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430

QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401

Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

R:SUZT 12  
A:G79546  
I: AAG79546 standard; protein; 484 AA.  
X: AAG79546;  
X: 10-DEC-2002 (first entry)  
D: Full length human TrpRS.  
C: Full length human TrpRS.  
T2; tryptophanyl-tRNA synthase; TrpRS; ocular neovascularisation;  
neovascular eye disease; age-related macular degeneration;  
ocular complication; diabetes; rubrotoxic glaucoma; retinopathy;  
prematurity; keratitis; ischaemic retinopathy; sickle cell;  
pathological myopia; ocular histoplasmosis; pterygia;  
punitate innerchorioidopathy; retinal degeneration; growth factor;  
vascularisation; vascular endothelial cell function; angiogenesis.

OS Homo sapiens.  
Key Location/Qualifiers  
FT 49..484  
FT /label= mini TrpRS  
FT /note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"  
FT 71..484  
FT /label= T1  
FT /note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"  
FT 94..484  
FT /label= T2  
FT /note= "Recombinant construct, has N-terminal Met and C-terminal KLAALHHHHHH"  
FT 170..173  
FT /label= Signature sequence  
FT Misc-difference 205..207  
FT /note= "May be replaced by Glu-Leu-Arg"  
FT 349..353  
FT /label= Signature sequence

WO200267970-A1.  
PD 06-SEP-2002.  
PF 22-FEB-2002; 2002WO-US005185.  
PR 23-FEB-2001; 2001US-0270951P.  
PP (SCRI ) SCRIPPS RES INST.  
PA Schimmel P, Wakasugi K, Friedlander M;  
PI  
XX WPI; 2002-698635/75.

PT New polypeptides derived from human tryptophanyl-tRNA synthase, useful for inhibiting ocular neovascularization in a patient, or for treating neovascular eye diseases, e.g. rubrotoxic glaucoma, retinopathy, keratitis, or pterygia.

PS Example 1; Fig 1; 83pp; English.

CC This sequence represents full length human tryptophanyl-tRNA synthase (TrpRS). A cleavage product, T2, of recombinant human TrpRS is water-soluble and comprises residues 94-471 of full length TrpRS. The water-soluble T2 polypeptide is useful for inhibiting ocular neovascularisation in a patient. This polypeptide is useful for treating neovascular eye diseases, e.g. age-related macular degeneration, ocular complications of diabetes, rubrotoxic glaucoma, retinopathy of prematurity, keratitis, ischaemic retinopathy (e.g. sickle cell), pathological myopia, ocular histoplasmosis, pterygia, or punitate innerchorioidopathy. This polypeptide is particularly useful for treating retinal degeneration to prevent the damaging effects of trophic and growth factors, and for promoting vascularisation to retard retinal degeneration by enhancing blood flow to cells. These are also useful for regulating vascular endothelial cell function, and in particular, for inhibiting angiogenesis

XX Sequence 484 AA;

Query Match 100.0%; Score 2116; DB 5; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.4e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNKGEDATEAEEDFVDPWTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQRP 60

Db 71 SNKGEDATEAEEDFVDPWTQTSAGKIDYDKLIVRFGSSKIDKELINRIERATGQRP 130

QY 61 FLRRGIFFSHRQMNQVLDAVENKKPFVLYTGRGPSSSEAMVGHLPFTKWLQDVNVP 120

Db 131 FLRRGIFFSHRQMNQVLDAVENKKPFVLYTGRGPSSSEAMVGHLPFTKWLQDVNVP 190

QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKN 180

Db 191 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTPIFSDLDYMGSSGFYKN 250

QY 181 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPPAIQAAPSFNSPQIFRDRDTIOCLIPCA 240

Db 251 VKIKQHVTFNQVKGIFGFTDSDCIGIKISPPAIQAAPSFNSPQIFRDRDTIOCLIPCA 310

QY 241 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 300

Db 311 IDQDPYFRMTDRVAPRIGYKPKALLHSTFPFALQGAQTKMSASDPNSSIFLTDTAKQIKT 370

QY 301 KVKHAFSGGRDTIEHRQFGNCVDVDSFWYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360

Db 371 KVKHAFSGGRDTIEHRQFGNCVDVDSFWYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430

QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401

Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

## RESULT 13

ABU72384  
ID ABU72384 standard; protein; 484 AA.  
AC  
AC ABU72384;  
XX  
XX 16-JUN-2003 (first entry)  
DT  
XX Human full length Tryptophanyl tRNA synthetase/His tag.

XX Human; tryptophanyl tRNA synthetase; TrpRS; angiogenesis; cytostatic;  
KW vulnary; Rossmann fold nucleotide binding domain; chemokine; ENAP II;  
KW vascular endothelial cell; solid tumour; myocardial infarction; enzyme;  
KW endothelial monocyte-activating polypeptide II; tumour metastasis;  
KW wound healing; dermal ulcer; endothelialisation; vascular graft surgery;  
KW abdominal wound; coronary bypass surgery; gene therapy.



XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PF US2002182666-A1.  
 XX PF 05-DBC-3C02.  
 XX PF 21-MAR-2001; 2001US-00813718.  
 XX PF 21-MAR-2001; 2001US-00813718.  
 XX PF (SCHL/) SCHIMMEL P.  
 XX PF (WAKA/) WAKASUGI K.  
 XX PF Schimmel P, Wakasugi K;  
 XX PF WPI; 2003-340974/32.  
 XX PF N-PSDB; ACA64105.  
 XX PF New truncated tryptophanyl-tRNA synthetase polypeptide comprising a  
 XX PF Rossmann fold nucleotide binding domain or having chemokine activity  
 XX PF useful for e.g. for regulating angiogenesis and for treating myocardial  
 XX PF infarction.  
 XX PF Example 1; Page 41-42; 91pp; English.  
 XX PF The invention relates to an isolated polypeptide comprising a truncated  
 XX PF tryptophanyl-tRNA synthetase (TrpRS) polypeptide comprising a Rossmann  
 XX PF fold nucleotide binding domain or having chemokine activity. The isolated  
 XX PF polypeptide is capable of regulating vascular endothelial cell function.  
 XX PF TrpRS has a C-terminal domain containing an EMAP II (endothelial monocyte  
 XX PF activating polypeptide II, a proinflammatory cytokine)-like domain and  
 XX PF is similar in sequence to TrpRS. Also included are a polynucleotide  
 XX PF encoding TrpRS (or a polynucleotide 95% similar to it) a TrpRS epitope,  
 XX PF 5' and 3' deletions of the TrpRS polynucleotide, a recombinant vector  
 XX PF comprising an isolated TrpRS nucleic acid, a recombinant host cell  
 XX PF containing the TrpRS nucleic acid and expressing TrpRS, an isolated anti-  
 XX PF TrpRS antibody, producing truncations of TrpRS by treating it with a  
 XX PF protease, and suppressing angiogenesis, solid tumours or a condition that  
 XX PF would benefit from decreased angiogenesis in a mammal by administering to  
 XX PF the mammal the composition comprising the TrpRS truncated protein. The  
 XX PF isolated polypeptide is useful for the preparation of a pharmaceutical  
 XX PF composition for transdermal, transmucosal, enteral or parenteral  
 XX PF administration. The truncated tRNA synthetase polypeptide is useful for  
 XX PF research, diagnostic, prognostic and therapeutic applications. The tRNA  
 XX PF synthetase are useful for regulating vascular endothelial cell function,  
 XX PF particularly for regulating angiogenesis, for treating myocardial  
 XX PF infarction and solid tumour, and for suppressing tumour metastasis.  
 XX PF Angiogenic tRNA synthetase polypeptides are useful as wound healing  
 XX PF agents or for treating full thickness wounds such as dermal ulcers, in  
 XX PF the promotion of endothelialisation in vascular graft surgery, in the  
 XX PF treatment of abdominal wounds where there is a high risk infection, in  
 XX PF conjunction with coronary bypass surgery by stimulating the growth of the  
 XX PF transplanted tissue, and in gene therapy. The present sequence is a Trp  
 XX PF tRNA synthetase protein (full length, truncated or mutant) with a His  
 XX PF affinity tag  
 XX PF Sequence 484 AA;  
 XX PF Query Match 100.0%; Score 2116; DB 6; Length 484;  
 XX PF Best Local Similarity 100.0%; Pred. No. 1.4e-209;  
 XX PF Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX PF 1 SNHGPDATERAEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINIRBRTGQRPHH 60  
 XX PF 71 SNHGPDATERAEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINIRBRTGQRPHH 130  
 XX PF 61 FLRRGIFFSHRDMMQVLDAVENKKPFYVYTGRRGSSSEAMHVGHILPFTKWLQDFNVVP 120  
 XX PF 131 FLRRGIFFSHRDMMQVLDAVENKKPFYVYTGRRGSSSEAMHVGHILPFTKWLQDFNVVP 190  
 XX PF 121 LVIQWTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFFSDDLDMGSSGYKN 180

Db 191 LVIQWTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFFSDDLDMGSSGYKN 250  
 QY 181 VKIQKHVTENQVKGIFGFTDSDICIGKISFPALQAAPSFNSPQIFDRDIDICLIPCA 240  
 Db 251 VKIQKHVTENQVKGIFGFTDSDICIGKISFPALQAAPSFNSPQIFDRDIDICLIPCA 310  
 QY 241 IDGQPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAOTKMSASDPNSSIFLTDPAKQIKT 300  
 Db 311 IDGQPYFRMTRDVAPRIGYKPKALLHSTFFPALQGAOTKMSASDPNSSIFLTDPAKQIKT 370  
 QY 301 KVNKHAFFSGRDTIEHRQFGNCDVDSVMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 Db 371 KVNKHAFFSGRDTIEHRQFGNCDVDSVMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430  
 QY 361 KKALIEVLQPLIAEHQARKEVTDIVKSPMTPRKLSFDFQ 401  
 Db 431 KKALIEVLQPLIAEHQARKEVTDIVKSPMTPRKLSFDFQ 471

## RESULT 14

ADE25762

ID ADE25762 standard; protein; 471 AA.

XX AC ADE25762;

XX DT 29-JAN-2004 (first entry)

XX DE Human protein differentially expressed in foam cells #39.

XX KW Human; differential expression; foam cell; LPS; lipopolysaccharide;

XX KW cardiovascular disease; atherosclerosis.

XX OS Homo sapiens.

XX PN US2003194721-A1.

XX PD 16-OCT-2003.

XX PF 18-SEP-2002; 2002US-00247671.

XX PR 19-SEP-2001; 2001US-0323784P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Mikita T, Shiffman D, Porter JG, Kaser MR;

XX DR WPI; 2003-875398/81.

XX DR N-PSDB; ADE25684.

XX CC Combination containing several polynucleotide that are differentially

XX CC expressed in foam cells and complements of the polynucleotides, useful

XX CC for diagnosing cardiovascular disease or atherosclerosis.

XX CC Discloure; SEQ ID NO 166; 37pp; English.

XX CC The invention relates to a combination comprising several polynucleotides

XX CC having any one of 127 sequences (S1) such as the sequence of human

XX CC calmodulin gene, human mRNA for KIA0930 protein, leukotriene A4

XX CC hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit

XX CC mRNA, etc., and their complements. The cDNAs are differentially expressed

XX CC in LPS (lipopolysaccharide)-treated foam cells. Also included are

XX CC obtaining an extended or full length gene from a library of nucleic acid

XX CC sequences, an expression vector containing the nucleic acids, a host cell

XX CC containing the vector, a purified polypeptide appearing as ADE25750 and

XX CC ADE25751, producing a protein by culturing the host cell, and a

XX CC composition comprising a purified antibody that specifically binds to the

XX CC proteins. The foam cell-expressed nucleic acids are useful for a high

XX CC throughput detection of differential expression of one or more

XX CC polynucleotides in a sample. The sample is from a subject with

XX CC atherosclerosis and comparison with a standard defines early, mid or late

XX CC stages of the disorder. The foam cell-expressed nucleic acids are useful

XX CC for high throughput screening of a library of molecules or compounds to

CC identify a ligand which binds a polynucleotide. The library is chosen  
 CC from DNA molecules, peptides, proteins and RNA molecules. The protein is  
 CC useful for a high throughput screening of library of molecules or  
 CC compounds to identify at least one ligand which specifically binds a  
 CC protein, for purifying a ligand from a sample for making an antibody. The  
 CC foam cell-expressed nucleic acids are useful for diagnosing  
 CC cardiovascular disorder. The foam cell-expressed nucleic acids are useful  
 CC as elements on a microarray which can be used for detecting related  
 CC polynucleotide in a sample, diagnosing cardiovascular disease,  
 CC atherosclerosis. The present sequence represents a protein differentially  
 CC expressed in JFS treated foam cells.

XX Sequence 471 AA;

Query Match 99.3%; Score 2101; DB 7; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 4.7e-208;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPWTQVTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 DB 71 SNHGDATEAEEDFVDPWTQVTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
 QY 61 FLRRGIFFSHRDMNQVLDAENKPPYLTYGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 120  
 DB 131 FLRRGIFFSHRDMNQVLDAENKPPYLTYGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 190  
 QY 121 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 DB 191 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250  
 QY 181 VVKIQKHVTENQVKGIFGFTSDDCIGKISFPALAHSTFFPALQGAQTKMSASDPNS 240  
 DB 251 VVKIQKHVTENQVKGIFGFTSDDCIGKISFPALAHSTFFPALQGAQTKMSASDPNS 310  
 QY 241 IDQDPYPRMTRDVAPRIGYKPKALHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 300  
 DB 311 IDQDPYPRMTRDVAPRIGYKPKALHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 370  
 QY 301 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 371 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
 QY 361 KKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSFDFQ 401  
 DB 431 KKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSFDFQ 471

RESULT 15  
 ADE76998  
 ID ADE76998 standard; protein; 471 AA.

XX AC ADE76998;

XX 29-JAN-2004 (first entry)

DE Human protein expressed in a liver disorder #46.

XX human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;  
 XX tumour; liver; inflammatory disorder; immune response disorder;  
 XX high-throughput screening; differential gene expression; gene therapy.

CS Homo sapiens.

IN US2003108871-A1.

XX 12-JUN-2003.

XX 30-JUL-2001; 2001US-00919039.

XX 28-JUL-2000; 2000US-0222113P.

XX (KASE/J) KASER M R.

PI Kaser MR;

XX WPI; 2004-031227/03.

DR N-PSDB; ADE76997.

XX Composition comprising several cDNAs that are differentially expressed in  
 PT treated human C3A liver cell cultures, useful for treating liver  
 PT disorders.

XX Claim 1; SEQ ID NO 163; 41pp; English.

CC The invention relates to a composition comprising several cDNAs that are  
 CC differentially expressed in a liver disorder. The composition is useful  
 CC for treating liver disorder such as hyperlipidaemia, hypertension, type  
 CC II diabetes, tumours of the liver and disorders of the inflammatory and  
 CC immune response. The composition is useful for a high-throughput method  
 CC of screening several molecules or compounds to identify a ligand which  
 CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a  
 CC high-throughput method for using a protein to screen several molecules or  
 CC compounds to identify at least one ligand which specifically binds the  
 CC protein which involves combining the protein encoded by the cDNA with  
 CC several of molecules or compounds under conditions to allow specific  
 CC binding, and detecting specific binding between the protein and a  
 CC molecule or compound, therefore identifying a ligand which specifically  
 CC binds the protein. The composition is useful for detecting and  
 CC quantifying differential gene expression, can be used in gene therapy, to  
 CC formulate prognosis and to design a treatment regimen and to monitor the  
 CC efficacy of treatment. The present sequence represents the amino acid  
 CC sequence of a protein encoded by a cDNA differentially expressed in a  
 CC liver disorder.

XX Sequence 471 AA;

Query Match 99.3%; Score 2101; DB 8; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 4.7e-208;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPWTQVTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRP 60  
 DB 71 SNHGDATEAEEDFVDPWTQVTSAGIDYDKLIVRFGSSKIDKELINRIERATGQRP 130  
 QY 61 FLRRGIFFSHRDMNQVLDAENKPPYLTYGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 120  
 DB 131 FLRRGIFFSHRDMNQVLDAENKPPYLTYGRGSPSEAMHVGHLIPFIETKWLQDVFNVP 190  
 QY 121 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 DB 191 LVIQMTDDEKYLWKDLTLDQAYDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250  
 QY 181 VVKIQKHVTENQVKGIFGFTSDDCIGKISFPALAHSTFFPALQGAQTKMSASDPNS 240  
 DB 251 VVKIQKHVTENQVKGIFGFTSDDCIGKISFPALAHSTFFPALQGAQTKMSASDPNS 310  
 QY 241 IDQDPYPRMTRDVAPRIGYKPKALHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 300  
 DB 311 IDQDPYPRMTRDVAPRIGYKPKALHSTFFPALQGAQTKMSASDPNSIFLTDTAQIKT 370  
 QY 301 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 360  
 DB 371 KVNKHAFGSGGRDTIEHRQFGNCVDVSVFMYLTFFLEDDDKLEQIRKDYTSGAMLTGEL 430  
 QY 361 KKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSFDFQ 401  
 DB 431 KKALIEVLQPLIAEHQARKEVTDIVKEFMTPRKLSFDFQ 471

Search completed: August 24, 2004, 17:50:32  
 Job time : 130 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 18:01:04 ; Search time 125 Seconds  
(without alignments)  
1008.130 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPWTV.....VTDEIVKEFTPRKLSDFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1295152 seqs, 31455058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| R.sult No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 2116  | 100.0       | 401    | 12    | US-10-080-839-13   |
| 2          | 2116  | 100.0       | 415    | 9     | US-09-813-718-14   |
| 3          | 2116  | 100.0       | 415    | 12    | US-10-080-839-5    |
| 4          | 2116  | 100.0       | 415    | 15    | US-10-240-532-14   |
| 5          | 2116  | 100.0       | 415    | 16    | US-10-240-527A-14  |
| 6          | 2116  | 100.0       | 437    | 9     | US-09-813-718-12   |
| 7          | 2116  | 100.0       | 437    | 12    | US-10-080-839-3    |
| 8          | 2116  | 100.0       | 437    | 15    | US-10-240-532-12   |
| 9          | 2116  | 100.0       | 437    | 16    | US-10-240-527A-12  |
| 10         | 2116  | 100.0       | 471    | 14    | US-10-126-467B-2   |
| 11         | 2116  | 100.0       | 471    | 15    | US-10-295-027-1234 |
| 12         | 2116  | 100.0       | 484    | 9     | US-09-813-718-10   |
| 13         | 2116  | 100.0       | 484    | 12    | US-10-080-839-1    |
| 14         | 2116  | 100.0       | 484    | 15    | US-10-240-532-10   |
| 15         | 2116  | 100.0       | 484    | 16    | US-10-240-527A-10  |

|    |        |      |     |    |                      |                    |
|----|--------|------|-----|----|----------------------|--------------------|
| 16 | 2101   | 99.3 | 471 | 10 | US-09-919-039-163    | Sequence 163, App  |
| 17 | 2101   | 99.3 | 471 | 14 | US-10-247-1671-166   | Sequence 166, App  |
| 18 | 2101   | 99.3 | 471 | 16 | US-10-408-765A-1235  | Sequence 1235, App |
| 19 | 2101   | 99.3 | 475 | 9  | US-09-925-302-558    | Sequence 558, App  |
| 20 | 2101   | 99.3 | 475 | 12 | US-09-925-302-558    | Sequence 558, App  |
| 21 | 1988   | 94.0 | 378 | 12 | US-10-080-839-12     | Sequence 12, App   |
| 22 | 1988   | 94.0 | 392 | 9  | US-09-813-718-16     | Sequence 16, App   |
| 23 | 1988   | 94.0 | 392 | 12 | US-10-080-839-7      | Sequence 7, Appl   |
| 24 | 1988   | 94.0 | 392 | 15 | US-10-240-532-16     | Sequence 16, Appl  |
| 25 | 1988   | 94.0 | 392 | 16 | US-10-240-527A-16    | Sequence 16, Appl  |
| 26 | 1938   | 91.6 | 475 | 14 | US-10-205-219-65     | Sequence 65, Appl  |
| 27 | 1326.5 | 62.7 | 408 | 16 | US-10-437-963-120838 | Sequence 120838,   |
| 28 | 1326.5 | 62.7 | 410 | 12 | US-10-424-599-146807 | Sequence 146807,   |
| 29 | 1302   | 61.5 | 429 | 12 | US-10-425-114-58867  | Sequence 58867, A  |
| 30 | 1200.5 | 56.7 | 424 | 14 | US-10-032-585-7632   | Sequence 7632, A   |
| 31 | 1125   | 53.2 | 433 | 14 | US-10-128-714-8545   | Sequence 8545, App |
| 32 | 1107.5 | 52.3 | 456 | 15 | US-10-320-797-3204   | Sequence 3204, App |
| 33 | 831    | 39.3 | 173 | 9  | US-09-925-302-855    | Sequence 855, App  |
| 34 | 831    | 39.3 | 173 | 12 | US-09-925-302-855    | Sequence 855, App  |
| 35 | 704.5  | 33.3 | 324 | 16 | US-10-437-963-108176 | Sequence 108176,   |
| 36 | 414.5  | 19.6 | 179 | 14 | US-10-128-714-3545   | Sequence 3545, App |
| 37 | 398    | 18.8 | 137 | 16 | US-10-437-963-145795 | Sequence 145795,   |
| 38 | 328    | 15.5 | 85  | 9  | US-09-813-718-45     | Sequence 45, Appl  |
| 39 | 328    | 15.5 | 85  | 15 | US-10-240-532-45     | Sequence 45, Appl  |
| 40 | 328    | 15.5 | 85  | 16 | US-10-240-527A-45    | Sequence 45, Appl  |
| 41 | 314    | 14.8 | 113 | 12 | US-10-424-599-146805 | Sequence 146805,   |
| 42 | 292    | 13.8 | 85  | 9  | US-09-813-718-46     | Sequence 46, Appl  |
| 43 | 292    | 13.8 | 85  | 15 | US-10-240-532-46     | Sequence 46, Appl  |
| 44 | 292    | 13.8 | 85  | 16 | US-10-240-527A-46    | Sequence 46, Appl  |
| 45 | 273.5  | 12.9 | 85  | 9  | US-09-813-718-48     | Sequence 48, Appl  |

#### ALIGNMENTS

##### RESULT 1

```
US-10-080-839-13
; Sequence 13, Application US/10080839
; Publication No. US20030017567A1
; GENERAL INFORMATION:
; APPLICANT: Wakasugi, Paul
; APPLICANT: Friedlander, Martin
; TITLE OF INVENTION: Tryptophanyl-tRNA Synthetase Derived
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Arginogenesis
; FILE REFERENCE: TSRI-813.1
; CURRENT APPLICATION NUMBER: US/10/060,839
; PRIOR FILING DATE: 2002-02-22
; PRIOR FILING DATE: 2002-02-22
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-839-13
```

|                       |                 |  |           |             |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match           | 100.0%          | Score 2116;  | DB 12;    | Length 401; |
| Best Local Similarity | 100.0%          | Pred. No. 3.5e-205;  |           |             |
| Matches 401;          | Conservative 0; | Mis-matches 0;   | Indels 0; | Gaps 0;     |
| Qy                    | 1               | SNHGPDATAEEDFVDPWTVTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHH  | 60        |             |
| Db                    | 1               | SNHGPDATAEEDFVDPWTVTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHH  | 60        |             |
| Qy                    | 61              | FLRGIFFSHRDMMQVLDAYENKPPFYLYTGRGSSSEAMHVGHILPFIPTFKWLQDVFNVP | 120       |             |
| Db                    | 61              | FLRGIFFSHRDMMQVLDAYENKPPFYLYTGRGSSSEAMHVGHILPFIPTFKWLQDVFNVP | 120       |             |
| Qy                    | 121             | LVIOMTDEKYLWDLTDQAYGDAVENAKDIIACGPDINKTFISDLDMGMSGGFYKN      | 180       |             |
| Db                    | 121             | LVIOMTDEKYLWDLTDQAYGDAVENAKDIIACGPDINKTFISDLDMGMSGGFYKN      | 180       |             |

181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240  
 181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240  
 241 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 300  
 241 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 300  
 301 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 301 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 361 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 401  
 361 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 401

## RESULT 2

US-09-813-718-14  
 ; Sequence 14, Application US/09813718  
 ; Publication No. US20020182666A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schimmel, Paul  
 ; APPLICANT: Wakasugi, Keisuke  
 ; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
 ; FILE REFERENCE: 00-221  
 ; CURRENT APPLICATION NUMBER: US/09/813,718  
 ; CURRENT FILING DATE: 2001-03-21  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 14  
 ; LENGTH: 415  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: human  
 ; OTHER INFORMATION: supermini Trprs in pET20B

## US-09-813-718-14

Query Match 100.0%; Score 2116; DB 9; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60  
 2 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61  
 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLTGRGSPSEAMHVGHLPFTTKWLQDVFNVP 120  
 62 FLRRGIFFSHRDMNQVLDAVENKKPFYLTGRGSPSEAMHVGHLPFTTKWLQDVFNVP 121  
 121 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 122 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240  
 182 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 241  
 241 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 300  
 242 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 301  
 301 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 302 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
 361 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 401  
 362 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 402

## RESULT 3

US-10-080-839-5  
 ; Sequence 5, Application US/10080339  
 ; Publication No. US20030017564A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schimmel, Paul  
 ; APPLICANT: Wakasugi, Keisuke  
 ; APPLICANT: Friedlander, Martin  
 ; TITLE OF INVENTION: Tryptophanyl-tRNA Synthetase Derived  
 ; FILE REFERENCE: TSRI-813.1  
 ; CURRENT APPLICATION NUMBER: US/10/080,839  
 ; CURRENT FILING DATE: 2002-02-22  
 ; PRIOR APPLICATION NUMBER: 60/270,951  
 ; PRIOR FILING DATE: 2001-02-23  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 415  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Cleavage Product T1 of recombinant human Trprs

## US-10-080-839-5

Query Match 100.0%; Score 2116; DB 12; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 60  
 2 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHH 61  
 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLTGRGSPSEAMHVGHLPFTTKWLQDVFNVP 120  
 62 FLRRGIFFSHRDMNQVLDAVENKKPFYLTGRGSPSEAMHVGHLPFTTKWLQDVFNVP 121  
 121 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
 122 LVIQMTDDEKYLWKDLTDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 181  
 181 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240  
 182 VVKIQHVTNFQVKGIFGFTSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 241  
 241 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 300  
 242 IDODPYFMTDRVAPRIGYKPKALLHSTFFPALQAGTQMSASDPSNSSFILDTAKQIKT 301  
 301 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360  
 302 KVNKHAFFSGGRDTIEHRQFGNCDDVVSFMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 361  
 361 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 401  
 362 KKALEVLQPLIAEHOARKEVTDEIVKEFMTPRKLSDFQ 402

## RESULT 4

US-10-240-532-14  
 ; Sequence 14, Application US/10240532  
 ; Publication No. US20040009163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schimmel, Paul  
 ; APPLICANT: Wakasugi, Keisuke  
 ; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
 ; FILE REFERENCE: TSRI 720.1  
 ; CURRENT APPLICATION NUMBER: US/10/240,532  
 ; CURRENT FILING DATE: 2002-03-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/08975  
 ; PRIOR APPLICATION NUMBER: 2001-03-21

PRIOR APPLICATION NUMBER: US 60/193,471  
PRIOR APPLICATION NUMBER: 2000-03-31  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 415  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human  
OTHER INFORMATION: supermini TrpRS in pET20B  
US-10-240-527A-14

Query Match  
Best Local Similarity 100.0%; Score 2116; DB 15; Length 415;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 61  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 62 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 121  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 180  
ID 122 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 181  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDRDTDIQCLIPCA 240  
ID 182 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDRDTDIQCLIPCA 241  
CY 241 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDCTAKQIKT 300  
ID 242 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDCTAKQIKT 301  
CY 301 KVNKHAFFSGRDTIEHRQFGNCVDVDFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
ID 362 KVNKHAFFSGRDTIEHRQFGNCVDVDFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 361  
CY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401  
ID 362 KKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 402

RESULT 5  
US-10-240-527A-14  
Sequence 14, Application US/10240527A  
Publication No. US20040152079A1  
GENERAL INFORMATION:  
APPLICANT: SCHIMMEL, Paul  
APPLICANT: WAKASUGI, Keisuke  
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase  
TITLE OF INVENTION: Polypeptides Useful for the Regulation of Angiogenesis  
FILE REFERENCE: TSRI 720.2  
CURRENT APPLICATION NUMBER: US/10/240,527A  
CURRENT FILING DATE: 2002-09-30  
PRIOR APPLICATION NUMBER: PCT/US01/08966  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/193,471  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 415  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Human supermini TrpRS in pET20B  
US-10-240-527A-14

Query Match  
Best Local Similarity 100.0%; Score 2116; DB 16; Length 415;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 3.7e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 2 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 61  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 62 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 121  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 180  
ID 122 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 181  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDRDTDIQCLIPCA 240  
ID 182 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDRDTDIQCLIPCA 241  
CY 241 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDCTAKQIKT 300  
ID 242 IDQDPYFRMTROVAPRIGYKPKALLHSTFFPALQGAQTKMSASDPNSIFLTDCTAKQIKT 301  
CY 301 KVNKHAFFSGRDTIEHRQFGNCVDVDFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
ID 362 KVNKHAFFSGRDTIEHRQFGNCVDVDFMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 361  
CY 361 KKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 401  
ID 362 KKALIEVLOPLIAEHQARRKEVTDEIVKEFMTPRKLSDFDQ 402

## RESULT 6

US-09-813-718-12  
Sequence 12, Application US/09813718  
Publication No. US20020182666A1  
GENERAL INFORMATION:  
APPLICANT: Schimmel, Paul  
APPLICANT: Wakasugi, Keisuke  
TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
TITLE OF INVENTION: The Regulation of Angiogenesis  
FILE REFERENCE: 00-221  
CURRENT APPLICATION NUMBER: US/09/813,718  
CURRENT FILING DATE: 2001-03-21  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human mini  
OTHER INFORMATION: TrpRS in pET20B  
US-09-813-718-12

Query Match  
Best Local Similarity 100.0%; Score 2116; DB 9; Length 437;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 60  
ID 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRFGSSKIDKELINRIERATGQRPHE 83  
CY 61 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 120  
ID 84 FLRRGIFFSHRDMNQVLDAYENKPFYLYTGRGSPSEAMHVGHLPFIPTKWLQDVFNVP 143  
CY 121 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 180  
ID 144 LVITQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTIFSDLOYMGSSGFYKN 203  
CY 181 VVKIQKHVTFFNOVKGIFGFTDSDCIGKISFPATQAAPSFNSPPOIFRDRDTDIQCLIPCA 240

Db 204 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

RESULT 7  
US-10-080-839-3  
; Sequence 3, Application US/10080839  
; Publication No. US20030017564A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; APPLICANT: Friedlander, Martin  
; TITLE OF INVENTION: tryptophanyl-tRNA Synthetase Derived  
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Angiogenesis  
; FILE REFERENCE: TSRI-813.1  
; CURRENT APPLICATION NUMBER: US/10/080,839  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: 60/270,951  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: human mini TrpS in pET20B

US-10-080-839-3  
Query Match 100.0%; Score 2116; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
Db 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
QY 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 120  
Db 84 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 240  
Db 204 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

US-10-080-839-3  
Query Match 100.0%; Score 2116; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
Db 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
QY 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 120  
Db 84 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 240  
Db 204 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

RESULT 8  
US-10-240-532-12  
; Sequence 12, Application US/10240532  
; Publication No. US20040009163A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For  
; TITLE OF INVENTION: The Regulation of Angiogenesis  
; FILE REFERENCE: TSRI 720.1  
; CURRENT APPLICATION NUMBER: US/10/240,532  
; CURRENT FILING DATE: 2002-03-30  
; PRIOR APPLICATION NUMBER: PCT/US01/08975  
; PRIOR APPLICATION NUMBER: 2001-03-21  
; PRIOR APPLICATION NUMBER: JS 60/193,471  
; PRIOR APPLICATION NUMBER: 2000-03-31  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: human mini  
; OTHER INFORMATION: TrpS in pET20B

US-10-240-532-12  
Query Match 100.0%; Score 2116; DB 15; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 60  
Db 24 SNHGPDATAEEDFVDPWTQTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRP 83  
QY 61 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 120  
Db 84 FLRRGIFFSHRDMNQVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFIFTKWLODVNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 144 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 240  
Db 204 VKIQKHTVFNQVKGIFGFTSDCIGKISFPALQAPSFNSFPQIPRDRDIOCLIPCA 263  
QY 241 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 300  
Db 264 IDQDPYFRMTRDVAPRIGYPKPKALLHSTFFPALQAGTKMSASDPNSSIFLTDATAKQIKT 323  
QY 301 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 360  
Db 324 KVNKHAFFSGGRDTEIEHRQFGNCDDVDVSPMYLTFPLEDDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
Db 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

RESULT 9  
US-10-240-527A-12  
; Sequence 12, Application US/10240527A  
; Publication No. US20040152079A1  
; GENERAL INFORMATION:  
; APPLICANT: Schimmel, Paul  
; APPLICANT: Wakasugi, Keisuke  
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase  
; TITLE OF INVENTION: Polypeptides Useful for the Regulation of Angiogenesis  
; FILE REFERENCE: TSRI 720.2  
; CURRENT APPLICATION NUMBER: US/10/240,527A  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: PCT/US01/08966

;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/193,471  
;; PRIOR FILING DATE: 2003-03-31  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: PastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 437  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Human mini TrpRS in pET20B  
U3-10-240-527A-12

Query Match 100.0%; Score 2116; DB 16; Length 437;  
Best Local Similarity 100.0%; Pred. No. 4e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q/ 1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP 60  
D 24 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP 83  
QY 61 FLRGIFFSHRDMQVLDAYENKPFYLYTGRGPSSEAMHVGHLIPFFITKWLQDVFNVP 120  
D 84 FLRGIFFSHRDMQVLDAYENKPFYLYTGRGPSSEAMHVGHLIPFFITKWLQDVFNVP 143  
QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
D 144 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 203  
QY 181 VKIKQHVTFNQVKGIFGFTSDCIGKISFPAIQAPSFNSFPQIFRDRDTIQCLIPCA 240  
D 204 VKIKQHVTFNQVKGIFGFTSDCIGKISFPAIQAPSFNSFPQIFRDRDTIQCLIPCA 263  
QY 241 IDQPYFMRTRDVAIRIGYPKPALESFFPALQAGTQMSASDPNSIFLDTAKQIKT 300  
D 264 IDQPYFMRTRDVAIRIGYPKPALESFFPALQAGTQMSASDPNSIFLDTAKQIKT 323  
QY 301 KVNKHAFGSGRDTIEHRQFGNCVDVVSFMYLTFPLEDDKLEQIRKDYTSGLMTGEL 360  
D 324 KVNKHAFGSGRDTIEHRQFGNCVDVVSFMYLTFPLEDDKLEQIRKDYTSGLMTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
D 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424

RESULT 10  
US-10-126-467B-2  
;; Sequence 2, Application US/10126467B  
;; Publication No. US20030059797A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Paley, Elena  
;; TITLE OF INVENTION: ANIMAL MODEL OF AND TEST FOR ALZHEIMER'S DISEASE  
;; FILE REFERENCE: PALL-111  
;; CURRENT APPLICATION NUMBER: US/10/126,467B  
;; CURRENT FILING DATE: 2002-11-19  
;; PRIOR APPLICATION NUMBER: 60/284,980  
;; PRIOR FILING DATE: 2001-04-19  
;; PRIOR APPLICATION NUMBER: 09/513,895  
;; PRIOR FILING DATE: 2000-02-28  
;; PRIOR APPLICATION NUMBER: 09/384,869  
;; PRIOR FILING DATE: 1999-08-27  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 471  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
U3-10-126-467B-2

Query Match 100.0%; Score 2116; DB 14; Length 471;  
Best Local Similarity 100.0%; Pred. No. 4.5e-205;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP 60  
D 71 SNHGPDTEAEEDFVDPMTVQTSSAKGIDYDKLIIVRFGSSKIDKELINRIERATGQRP 130  
QY 61 FLRGIFFSHRDMQVLDAYENKPFYLYTGRGPSSEAMHVGHLIPFFITKWLQDVFNVP 120  
D 131 FLRGIFFSHRDMQVLDAYENKPFYLYTGRGPSSEAMHVGHLIPFFITKWLQDVFNVP 190  
QY 121 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
D 131 LVIQMTDDEKYLWKDLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250  
QY 181 VKIKQHVTFNQVKGIFGFTSDCIGKISFPAIQAPSFNSFPQIFRDRDTIQCLIPCA 240  
D 251 VKIKQHVTFNQVKGIFGFTSDCIGKISFPAIQAPSFNSFPQIFRDRDTIQCLIPCA 310  
QY 241 IDQPYFMRTRDVAIRIGYPKPALESFFPALQAGTQMSASDPNSIFLDTAKQIKT 300  
D 311 IDQPYFMRTRDVAIRIGYPKPALESFFPALQAGTQMSASDPNSIFLDTAKQIKT 370  
QY 301 KVNKHAFGSGRDTIEHRQFGNCVDVVSFMYLTFPLEDDKLEQIRKDYTSGLMTGEL 360  
D 371 KVNKHAFGSGRDTIEHRQFGNCVDVVSFMYLTFPLEDDKLEQIRKDYTSGLMTGEL 430  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401  
D 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 471

RESULT 11  
US-10-295-027-1234  
;; Sequence 1234, Application US/10295027  
;; Publication No. US20030232350A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Afar, Daniel  
;; APPLICANT: Aziz, Natasha  
;; APPLICANT: Ginsberg, Wendy M.  
;; APPLICANT: Gish, Kurt C.  
;; APPLICANT: Glynn, Richard  
;; APPLICANT: Hevezi, Peter A.  
;; APPLICANT: Mack, David H.  
;; APPLICANT: Murray, Richard  
;; APPLICANT: Watson, Susan R.  
;; APPLICANT: Eos Biotechnology, Inc.  
;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
;; FILE REFERENCE: 018501-012500US  
;; CURRENT APPLICATION NUMBER: US/10/295,027  
;; CURRENT FILING DATE: 2002-11-13  
;; PRIOR APPLICATION NUMBER: US 09/663,733  
;; PRIOR FILING DATE: 2000-09-15  
;; PRIOR APPLICATION NUMBER: US 60/350,666  
;; PRIOR FILING DATE: 2001-11-13  
;; PRIOR APPLICATION NUMBER: US 60/335,394  
;; PRIOR FILING DATE: 2001-11-15  
;; PRIOR APPLICATION NUMBER: US 60/332,464  
;; PRIOR FILING DATE: 2001-11-21  
;; PRIOR APPLICATION NUMBER: US 60/334,393  
;; PRIOR FILING DATE: 2001-11-29  
;; PRIOR APPLICATION NUMBER: US 60/340,376  
;; PRIOR FILING DATE: 2001-12-14  
;; PRIOR APPLICATION NUMBER: US 60/347,211  
;; PRIOR FILING DATE: 2002-01-08  
;; PRIOR APPLICATION NUMBER: US 60/347,349  
;; PRIOR FILING DATE: 2002-01-10  
;; PRIOR APPLICATION NUMBER: US 60/355,250  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: US 60/356,714  
;; PRIOR FILING DATE: 2002-02-13  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 1386  
;; SOFTWARE: PatentIn Ver. 2.1

```
/ SEQ ID NO 1234
/ LENGTH: 471
/ TYPE: PRT
/ ORGANISM: Homo sapiens
U:-10-295-027-1234

Query Match      100.0%; Score 2116; DB 15; Length 471;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
D: 71 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 120
D: 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 190
QY 121 LV1QMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFISDLYMGSSGGFYKN 180
D: 191 LV1QMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFISDLYMGSSGGFYKN 250
QY 181 VVKIQKHVTFNQVKGIFGFTSDSCIGIKISPPALQAAPSFNSFPQIFRDRDIIQCLIPCA 240
D: 251 VVKIQKHVTFNQVKGIFGFTSDSCIGIKISPPALQAAPSFNSFPQIFRDRDIIQCLIPCA 310
QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSASDPNSSIFLTDATAKQIKT 300
D: 311 IDQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSASDPNSSIFLTDATAKQIKT 370
QY 301 KVNKHAFFSGGRDTEIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 360
D: 371 KVNKHAFFSGGRDTEIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 430
QY 361 KXALIEVLQPLIAEHQARRKEVTDEIVKEPMTPRKLSDFDQ 401
D: 431 KXALIEVLQPLIAEHQARRKEVTDEIVKEPMTPRKLSDFDQ 471

RESULT 13
US-10-080-839-1
; Sequence 1, Application US/10080839
; Publication No. US20030017564A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakaugi, Keisuke
; APPLICANT: Friedlander, Martin
; TITLE OF INVENTION: Tryptophanyl-tRNA Synthetase Derived
; TITLE OF INVENTION: Polypeptides Useful For The Regulation of Arg-oogenesis
; FILE REFERENCE: TSRI-813.1
; CURRENT APPLICATION NUMBER: US/10/080,839
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,951
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant human tprRS
US-10-080-839-1

Query Match      100.0%; Score 2116; DB 12; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
D: 71 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 120
D: 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 190
QY 121 LV1QMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFISDLYMGSSGGFYKN 180
D: 191 LV1QMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGPDINKTFISDLYMGSSGGFYKN 250
QY 181 VVKIQKHVTFNQVKGIFGFTSDSCIGIKISPPALQAAPSFNSFPQIFRDRDIIQCLIPCA 240
D: 251 VVKIQKHVTFNQVKGIFGFTSDSCIGIKISPPALQAAPSFNSFPQIFRDRDIIQCLIPCA 310
QY 241 IDQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSASDPNSSIFLTDATAKQIKT 300
D: 311 IDQDPYFRMTRDVAPRIGYKPKALLHSTFPALQGAOTKMSASDPNSSIFLTDATAKQIKT 370
QY 301 KVNKHAFFSGGRDTEIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 360
D: 371 KVNKHAFFSGGRDTEIEHRQFGGNCVDVSMYLTFFLEDDDKLEIQIRKDYTSGLMTGEL 430

RESULT 12
US-09-813-718-10
; Sequence 10, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakaugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 90-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human
; OTHER INFORMATION: full-length TprRS in pET20B
US-09-813-718-10

Query Match      100.0%; Score 2116; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 60
D: 71 SNHGPDATAEEDFVDPWTQTSAGIDYDKLIVRGSSKIDKELINRIERATGQRP 130
QY 61 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 120
D: 131 FLRRGIFFSHRDNNQVLDAYENKKPFYLYTGRGPSSEAMVGHLLPIFTKWLQDVNP 190
```



```

Db 371 KVNKAFSGGRDTIEHRQFGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

RESULT 14
US-10-240-532-10
; Sequence 10, Application US/10240532
; Publication No. US20040009163A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; FILE REFERENCE: TSRI 720.1
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/10/240,532
; PRIOR APPLICATION NUMBER: PCT/US01/08975
; PRIOR APPLICATION NUMBER: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/193,471
; PRIOR APPLICATION NUMBER: 2000-03-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: human
; OTHER INFORMATION: full-length TrpRS in pET20B
US-10-240-532-10

Query Match 100.0%; Score 2116; DB 15; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 60
Db 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 130
QY 61 FLRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMHGVLIPFTTKWLQDVNP 120
Db 131 FLRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMHGVLIPFTTKWLQDVNP 190
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
Db 191 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
QY 181 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240
Db 251 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 310
QY 241 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 300
Db 311 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 370
QY 301 KVNKAFSGGRDTIEHRQFGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
Db 371 KVNKAFSGGRDTIEHRQFGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

```

```

RESULT 15
US-10-240-527A-10
; Sequence 10, Application US/10240527A
; Publication No. US20040152079A1
; GENERAL INFORMATION:
; APPLICANT: SCHIMMEL, Paul

```

```

; APPLICANT: WAKASUGI, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase
; FILE REFERENCE: Polypeptides Useful for the Regulation of Angiogenesis
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/10/240,527A
; PRIOR APPLICATION NUMBER: PCT/US01/08966
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/193,471
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Versiocr 4.0
; SEQ ID NO 10
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Human full-length TrpRS in pET20B
US-10-240-527A-10

Query Match 100.0%; Score 2116; DB 16; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-205;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 60
Db 71 SNHGDATEAEEDFVDPMTVQTSSAKGIDYDKLIIVFGSSKIDKELINRIERATGQRP 130
QY 61 FLRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMHGVLIPFTTKWLQDVNP 120
Db 131 FLRGIFFSHRDMQVLDAYENKKEFYLYTGRGSPSEAMHGVLIPFTTKWLQDVNP 190
QY 121 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180
Db 191 LVIQMTDDEKYLWKDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250
QY 181 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 240
Db 251 VVKIQKHVTENQKGI FGFTSDSDCIKISFPALQAPSFNSFPQIFRDRDTIQCLIPCA 310
QY 241 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 300
Db 311 IDQDPYFRMTRDVAPRIGYKPKPALLHSTFPALQAGATKMSASDPNSSIFLDTAKQIKT 370
QY 301 KVNKAFSGGRDTIEHRQFGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 360
Db 371 KVNKAFSGGRDTIEHRQFGNCVDVSPMYLTFFLEDDDKLEQIRKDYTSGLMTGEL 430
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 401
Db 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDQ 471

```

Search completed: August 24, 2004, 19:21:54  
Job time : 127 secs

